Wirtschafts informatik 2017

TAGUNGSBAND

13. Internationale Tagung Wirtschaftsinformatik

„Towards Thought Leadership in Digital Transformation"

12. - 15. Februar 2017, Universität St.Gallen

Jan Marco Leimeister und Walter Brenner (Hrsg.)

Institut für Wirtschaftsinformatik
Universität St.Gallen

Jan Marco Leimeister und Walter Brenner (Hrsg.)

Towards Thought Leadership in Digital Transformation

13. Internationale Tagung Wirtschaftsinformatik (WI 2017)

Tagungsband

Wirtschaftsinformatik 2017

Institut für Wirtschaftsinformatik

Universität St.Gallen

Vorwort der Tagungsleitung

Die Internationale Tagung Wirtschaftsinformatik ist die Hauptkonferenz der deutschsprachigen Wirtschaftsinformatik Community und war mit der 13. Ausgabe unter dem Konferenzmotto „Towards Thought Leadership in Digital Transformation“ in St. Gallen ein ganz besonderes Austauschforum zwischen Wissenschaft und Praxis für innovative und relevante Themen der Wirtschaftsinformatik und richtete sich an Wissenschaftler und Studierende mit Interesse an der Forschung und insbesondere Praktiker mit Bezug zur Wirtschaftsinformatik.

Es war uns als Gastgeber des Instituts für Wirtschaftsinformatik (IWI-HSG) der Universität St.Gallen und den Lehrstühlen von Prof. Dr. Jan Marco Leimeister und Prof. Dr. Walter Brenner eine Ehre die WI 2017 in St.Gallen ausrichten zu dürfen.

Es war uns zudem eine grosse Freude mit über 750 Teilnehmenden aus Wissenschaft und Praxis und der deutschsprachigen Wirtschaftsinformatik-Community die WI 2017 zu bestreiten. Wir hoffen die Tage in St. Gallen werden allen Teilnehmenden und Mitwirkenden in guter und bleibender Erinnerung bleiben.

In 14 wissenschaftlichen Tracks, einem Start-Up Track und dem Student Track bot die Konferenz vielfältige Möglichkeiten für einen intensiven Austausch zwischen Wissenschaft und Praxis. Besonders hervorheben möchten wir die interessanten Präsentationen im Rahmen des CIO Dialogs, das Doctoral und Junior Faculty Consortium sowie die Prototype & Experiment Session. Zusätzlich bot das Format des Prototype & Experiment Tracks mit einem neuen Einreichungsformat die Möglichkeit, innovative und gestaltungsorientierte Forschung im Rahmen der Konferenz interaktiv vor einem grossen Publikum zu präsentieren.

Weiterhin haben wir vor der Konferenz bei der WI 2017 einen neuartigen und innovativen Review-Prozess eingeführt, um die Geschwindigkeit und auch Qualität der Reviews für Autoren deutlich zu erhöhen. Hierzu haben, unter Leitung der Conference Chairs und des Reviewmanagements, die Track Chairs mit Hilfe Ihrer Associate Editor einen bisher einzigartigen Double-blind Peer Review Prozess durchgeführt. Dieser neuartige Prozess hat dabei den über 200 Associate Editors, welche sich im Vorfeld für die Tracks als Editors bewerben konnten, eine besonders wichtige Rolle zugestanden. Im Rahmen des „Fast and Constructive Associate Editor Rejects“ wurde die Geschwindigkeit des Review-Prozesses stark erhöht und Last von den Schultern der über 600 Reviewer genommen. Die Associate Editors und Reviewer haben dann dafür Sorge getragen, dass Beiträge mindestens dreimal begutachtet wurden und Autoren ein reichhaltiges und konstruktives Feedback auf die Forschungsarbeiten erhalten haben. Damit möchten wir den Track Chairs, Associate Editors als auch Reviewern für die Unterstützung im ganzen Prozess unseren herzlichen Dank ausdrücken und die extrem wertvolle Arbeit der Wirtschaftsinformatik-Community hervorheben.

Insgesamt wurden auf der Konferenz 119 Beiträge präsentiert und in diesem Tagungsband abgedruckt, wobei die Annahmequote der wissenschaftlichen Vollbeiträge bei 32,3% liegt. Wissenschaftliche Beiträge konnten bei der WI 2017 zudem auch als Extended Abstract veröffentlicht werden, insbesondere um den Autoren die Möglichkeit zu geben, Forschungsarbeiten für die Publikation in Journals weiter aufzubereiten. In diesem Kontext möchten wir hier nochmals hervorheben, dass Autoren im Zuge der WI 2017 die Chance hatten, im Review-Prozess für eine Fast-Track Publikation in renommierte Journals (BISE, EM, EMISA, HMD, i-com, ISeB) der Wirtschaftsinformatik eingeladen zu werden.

Der vorliegende Tagungsband der WI 2017 liefert auf dieser Basis einen Überblick über innovative und relevante Themen sowie Herausforderungen der Wirtschaftsinformatik, welche unter dem Konferenzmotto „Towards Thought Leadership in Digital Transformation“ standen.

Unser Dank gilt dabei den zahlreichen Sponsoren, Partnern und Förderern, welche durch die substanzielle Unterstützung diesen Event in dieser Form erst möglich gemacht haben. Weiterhin möchten wir dem ganzen Organisationsteam sowie Dr. Christoph Peters als Projektleiter herzlich danken.

Abschliessend danken wir allen Teilnehmenden, dass Sie den Weg in die Schweiz nach St. Gallen gefunden haben und hoffen, dass Sie ein attraktives Programm mit zahlreichen Forschungsbeiträgen, Keynotes, Panels, Workshops aber auch Rahmenprogramm wahrnehmen konnten.

Ihre Conference Chairs,

Jan Marco Leimeister & Walter Brenner

Organisation der WI 2017

Konferenzverantwortliche

Prof. Dr. Jan Marco Leimeister
Conference Chair

Prof. Dr. Walter Brenner
Conference Chair

Dr. Christoph Peters
Projektleiter

Host - Institut für Wirtschaftsinformatik (IWI-HSG)

Prof. Dr. Andrea Back (Coordination Social Media)
Prof. Dr. Reinhard Jung (Coordination Workshops)
Prof. Dr. Robert Winter (Coordination Doctorial Consortium)

Kernteam der WI 2017

Susanne Gmünder
Organization Management and Support

Barbara Rohner
Sponsor Acquisition and Management

Dr. Jochen Müller
CIO-Dialog Management

Wissenschaftliche Mitarbeitende (Alphabetisch)

Christian Dremel
Conference Dinner Organization

Dr. Philipp Ebel
Academic Networking Dinner Organization

Philipp Haas
Video Production

Markus Handke
IT-Management

Andreas Janson
Review and Program Management

Nicolas Knop
Pub Crawl and Volunteer Organization

Sarah Oeste-Reiß
Communication Management

Sandro Storchenegger
IT Support

Studentische Helfer (Alphabetisch)

Benjamin Emmenegger
Conference App Development
Marvin Herbring
Corporate Design
Samir Ibrahim
Statement Video
Maik Madörin
Drone Footage
Dominik Schirle
Event Production
Fabrizia Vollenweider
Statement Video
Benedikt Wakonig
Real-Time Infoscreen

Erweitertes Team der WI 2017

Wissenschaftliche Mitarbeitende (Alphabetisch)

Vanessa Guggisberg (Social Media) | **Manuel Holler** (Company Workshop Support) | Stefan Kleinschmidt (Track Chair Meeting Organization) | **Volkmar Mrass** (Press Management) | **Jan Martin Persch** (Rating Tool Development) | **Prof. Dr. Matthias Söllner** (Digital Collaboration Challenge and Junior Faculty Consortium)

Studentische Helfer (Alphabetisch)

Kira Fastner (Conference Proceedings and Program Support) | **Andreas Günther** (Track Videos) | **Ha-nl Kwon** (Volunteer and Non-Cash Sponsor Acquisition)

Operative Unterstützung während der Konferenz

Prof. Dr. Stephan Aier I Prof. Dr. Antonia Albani I Nihan Altug I Alina Asisof I Zana Bajrami I Sandro Balduini I Simon Baumgartner I Jannis Beese I Sabine Berghaus I Michael Blaschke I Prof. Dr. Ivo Blohm I Dr. Ulrich Bretschneider I Maximilian Brosius I Victoria Bucher I Dominik Dellermann I David Durward I Sissy-Josefina Ernst I Daniel Gerster I Martina Hagenbuch I Prof. Dr. Kazem Haki I Jennifer Hehn I Manuel Holler I Alexandra Kessler I Robin Knote I Prof. Dr. Christiane Lehrer I Niklas Leicht I Mahei Li I Nikolaus Lipusch I Sandra Lugonjic I Nicoline Lührs I Tobias Mirsch I Gerard Neiditsch I Marcel Rhyn I Roman Rietsche I Prof. Dr. Peter Rohner I Raphael Schilling I Barbara Schmidt I Sofia Schöbel I Daniel Siegrist I Benedikt Simmert I Benjamin Spottke I Emanuel Stöckli I Julia Troll I Prof. Dr. Falk Übernickel I Marc von Niederhäuser I Alexander Wieneke I Rainer Winkler I Prof. Dr. Jochen Wulf

Sponsoren der WI 2017

Track Chairs

Becker, Jörg

Benlian, Alexander

Beverungen, Daniel

Böhmann, Tilo

Dinter, Barbara

Fleisch, Elgar

Hess, Thomas

Koch, Michael

Krämer, Jan

Krasnova, Hanna

Lechner, Ulrike

Legner, Christine

Leist, Susanne

Loos, Peter

Mädche, Alexander

Maier, Roland

Mendling, Jan

Möslein, Kathrin

Niehaves, Björn

Otto, Boris

Schumann, Matthias

Strahringer, Susanne

Strecker, Stefan

Sunyaev, Ali

Veit, Daniel

vom Brocke, Jan

Weitzel, Tim

Widjaja, Thomas

Associate Editors aller Tracks

Frederik Ahlemann | Stephan Aier | Paul Alpar | Rainer Alt | Jörn Altmann | Jürgen Anke | Sören Auer | Henning Baars | Christine Bauer | Sabine Baumann | Ulrike Baumöl | Daniel Beimborn | Martin Bichler | Markus Bick | Eva Bittner | Ivo Blohm | Freimut Bodendorf | Andreas Bogner | Tobias Brandt | Christoph Breidbach | Andreas Breiter | Michael H. Breitner | Ulrich Bretschneider | Angelika C. Bullinger-Hoffmann | Ralf Buschermöhle | Ricardo Büttner | Peter Buxmann | Michel Clement | Patrick Delfmann | Regine Dörbecker | Verena Dorner | Paul Drews | Andreas Eckhardt | Jan Fabian Ehmke | Stefan Eicker | Tatiana Ermakova | Werner Esswein | Benjamin Fabian | Carsten Felden | Peter Fettke | Stefan Feuerriegel | Kathrin Figl | Kerstin Fink | Kai Fischbach | Christoph Flath | Gilbert Fridgen | Burkhardt Funk | Daniel Fürstenau | Michaela Geierhos | Jin Gerlach | Martin Gersch | Heiko Gewald | Andrea Giessmann | Henner Gimpel | Peter Gluchowski | Michael Granitzer | Jens Grossklags | Jens Gulden | Liane Haak | Joerg Haake | Maik Hammerschmidt | Irina Heimbach | Gundula Heinatz Bürki | Peter Heinrich | Bernd Hellingrath | Dominik Herrmann | Georg Herzwurm | Andreas Hilbert | Hartmut Hoehle | Sara Hofmann | Uwe Hoppe | Thomas Huber | Heinrich Hussmann | Christian Janiesch | Monique Janneck | Florian Johannsen | Thomas John | Julia Jonas | Reinhard Jung | Dimitris Karagiannis | Hans-Georg Kemper | Doğan Kesdoğan | Stefan Kirn | Stefan Klein | Mathias Klier | Ralf Knackstedt | Bart Knijnenburg | Matthias Knoll | Alexander Kock | Axel Korthaus | Agnes Koschmider | Tobias Kowatsch | Oliver Krancher | Johann Kranz | Artus Krohn-Grimberghe | Thomas Kude | Karl Lang | Sven Laumer | Ulrike Lechner | Franz Lehner | Henrik Leopold | Stefan Lessmann | Christian Leyh | André Loske | Carsten Lucke | Christopher Lueg | Christian Maier | Rainer Malaka | Attila Marton | Jorge Marx Gómez | Sabine Matook | Christian Matt | Florian Matthes | Martin Matzner | Florian Michahelles | Mirjam Minor | Lars Mönch | Stefan Morana | Michael zur Muehlen | Benjamin Müller | Oliver Müller | Günter Müller | Claudia Müller-Birn | Jan Muntermann | Dirk Neumann | Peter Niemeyer | Verena Nitsch | Andreas Oberweis | Sven Overhage | Jan Pawlowski | Christoph Peters | Jella Pfeiffer | Alexander Pflaum | Irena Pletikosa | Jens Pöppelbuß | Wolfgang Prinz | Henderik Proper | Ulrich Remus | Christian Reuter | Alexander Richter | Rene Riedl | Bodo Rieger | Stefanie Rinderle-Ma | Susanne Robra | Maximilian Röglinger | Michael Rosemann | Christoph Rosenkranz | Angela Roth | Gerhard Satzger | Silvia Schacht | Daniel Schlagwein | Christian Schlereth | Rainer Schmidt | Albrecht Schmidt | Michael Scholz | Guido Schryen | Monika Schuhmacher | Carsten Schultz | Gerhard Schwabe | Thomas Setzer | Mario Silic | Elmar J. Sinz | Stefan Smolnik | Mariya Sodenkamp | Matthias Söllner | Martin Spann | Kai Spohrer | Dirk Stelzer | Stefan Stieglitz | Steffen Stock | Steffen Strassburger | Stefan Strohmeier | Leena Suhl | Ioannis Tarnanas | Frank Teuteberg | Stefan Thalmann | Frédéric Thiesse | Oliver Thomas | Ingo J. Timm | Manuel Trenz | Klaus Turowski | Virpi Kristiina Tuunainen | Falk Uebernickel | Nils Urbach | Jan vom Brocke | Heinz-Theo Wagner | Marcus Wagner | Kristin Weber | Anette Weisbecker | Helena Wenninger | Dirk Werth | Lauri Wessel | Markus Westner | Martin Wiener | Manuel Wiesche | Sue Williams | Axel Winkelmann | Till Winkler | Thomas Wöhner | Felix Wortmann | Jochen Wulf | Volker Wulf | Nancy Wünderlich | Rüdiger Zarnekow | Steffen Zimmermann | Andreas Zolnowski | Christoph Zott

Liste der Reviewer

Konstantin Aal | Puja Abbassi | Olga Abramova | Safa'a AbuJarour | Abiodun Adegbile | Michael Adelmeyer | Frederik Ahlemann | Simon Albrecht | Simon Alfano | Paul Alpar | Rainer Alt | Jörn Altmann | Jürgen Anke | Armin Anzenbacher | Emilio Arruda | Colin Atkinsons | Sören Auer | Dominik Augenstein | Okan Aydingül | Banu Aysolmaz | Henning Baars | Christian Baccarella | Vanessa Bachmaier | Filipe Barata | Dirk Basten | Christine Bauer | Sabine Baumann | Annika Baumann | Jörg Becker | Jannis Beese | Daniel Beimborn | Mariem Ben Rehouma | Benedict Bender | Axel Benjamins | Frank Bensberg | Peder Bergan | Benedikt Berger | Sabine Berghaus | Sören Bergmann | Ralph Bergmann | Lisa Berkemeier | Thomas Berlage | Jan Ole Berndt | Max Bernhagen | Hans Betke | Jan Betzing | Saskia Bick | Pius Bienz | Marija Bijekovic | Dominik Bilgeri | Eva Bittner | Michael Blaschke | Katja Bley | Carolin Bock | Alexander Bock | Martin Böckle | Julian KH Boha | Jens Borchers | Dominik Bork | Paul Bossauer | Sascha Bosse | Hendrik Brakemeier | Hans Brandt-Pook | Daniel Braunnagel | Michael Breidung | Michael H. Breitner | Christian Brenig | Uwe Bretschneider | Marcel Breuer | Jan Broer | Constantin Brosda | Maximilian Brosius | Erik Buchmann | Arne Buchwald | Christoph Buck | Ricardo Buettner | Daniel Burda | Alexa Burmester | Ralf Buschermöhle | Paris Buttfield-Addison | Ricardo Büttner | Peter Chamoni | Tommy Chan | Mathieu Chanson | Kevin Clees | Michel Clement | Catherine Cleophas | Nico Clever | Christian Comberg | Raffaele Conforti | Ann-Kristin Cordes | Carl Corea | André Dahlinger | Frank Danzinger | John D'Arcy | Gábor Darvasi | Stephan Daurer | Johannes Daxenberger | Mark de Reuver | Gert-Jan de Vreede | Claudia de Witt | Patrick Delfmann | Dominik Dellermann | Lukas Demetz | Birgit Demuth | Thomas Diefenbach | Phil Diegmann | Remco Dijkman | Georg Disterer | Leonhard Dobusch | Mateusz Dolata | Hüseyin Doluca | Regine Dörbecker | Verena Dorner | Andreas Drechsler | Tim Dreesen | Christian Dremel | Paul Drews | Sebastian Duerr | Nico Ebert | Katharina Ebner | Frank Effenberger | Jan Fabian Ehmke | Tom Eich | Dennis Eilers | Maik Eisenbeiß | Kathrin Eismann | Leonie Ellis | Adrian Engelbrecht | Oliver Englisch | Martin Englster | Julian Ereth | Tatiana Ermakova | Claus-Peter Ernst | Werner Esswein | Benjamin Fabian | Thomas Farrenkopf | Roland Fassauer | Maria Fay | Walid Fdhila | Thomas Fehlmann | Carsten Felden | Niclas Feldkamp | Michael Fellmann | Peter Fettke | Stefan Feuerriegel | Sven Feurer | Kathrin Figl | Hans-Georg Fill | Matthias Finck | Kai Fischbach | Thomas Fischer | Daniel Fischer | Christoph Flath | Stefan Fleischer | Alexander Fliaster | Sebastian Floerecke | Barbara Flügge | Jens Förderer | Giancarlo Fortino | Ulrik Franke | Alexander Frey | Gilbert Fridgen | Thomas Friedrich | Veronika Fries | Sascha Friesike | Hansjörg Fromm | Nora Fteimi | Daniel Fürstenau | Bernhard Gahr | Manfred Geiger | Christoph Gerber | Martin Gersch | Heiko Gewald | Andrea Giessmann | Niklas Goby | Matthias Goeken | Paul Grefen | Robert Gregory | Katharina Greve | Norbert Gronau | Birgit Großer | Jens Grossklags | Christian Grotherr | Daniel Guhl | Jens Gulden | Gunther Gust | Steffi Haag | Liane Haak | Phillip Haake | Francois Habryn | Irit Hadar | Martin Haferkorn | Ilse Hagerer | Christopher Hahn | Axel Hahn | Kazem Haki | Tom Hänel | Andre Hanelt | Jannis Hanke | Nicolai Hanner | Kjeld Hansen | David Harborth | Catherine Hardy | Anuja Hariharan | Stefan Willi Hart | Stefan Hattula | Johannes Haupt | Clemens Haussmann | Jonas Hedman | Margareta Heidt | Irina Heimbach | Robert Heininger | Peter Heinrich | Bernd Heinrich | Andreas Helferich | Bernd Hellingrath | Patrick Helmholz | Raoul Hentschel | Dominik Herrmann | Sven Herzberg | Thomas Hess | Erich Heumüller | Andreas Heusler | Knut Hinkelmann | Johannes Hintsch | Sebastian Hobert | Tore Hoel | David Hoffmann | Benjamin Hoffmann | Sara Hofmann | Friedrich Holotiuk | Katharina Hölzle | Konstantin Hopf | Uwe Hoppe | Stijn Hoppenbrouwers | Bettina Horlach | Dennis Horstkemper | Constantin Houy | Yanni Hu | Thomas

Huber | Thomas Hurni | Mirka Hyvärinen | Ilona Ilvonen | Florian Imgrund | Conrad Indiono | Nihal Islam | Nihal Islam | Olga Ivanova | Björn Ivens | Naoum Jamous | Christian Janiesch | Monique Janneck | Slinger Jansen | Andreas Janson | Andrija Javor | Christian Jentsch | Stefan Jesse | Denis Johann | Florian Johannsen | Thomas John | Julia Jonas | Dominik Jung | Johannes Jurgovsky | Monika Kaczmarek | Cristina Kadar | Marcus Kaiser | Andre Kaiser | Chinmay Kakatkar | Andreas Kaltenbrunner | Jermain Kaminski | Oliver Kasten | Soeren Kerner | Doğan Kesdoğan | Janina Kettenbohrer | Murat Kezer | Ateeq Khan | Mehmet Kilic | Thomas Kirks | Hans-Bernd Kittlaus | Michael Klafft | Janne Kleinhans | Stefan Kleinschmidt | Michael Klesel | Mathias Klier | Natalia Kliewer | Ralf Klischewski | Christoph Klötzer | Ralf Knackstedt | Bart Knijnenburg | Robin Knote | David Knuplesch | Achim Koberstein | Jörn Kobus | Alexander Kock | Elisa Konya-Baumbach | Somayeh Koohborfardhaghighi | Derck Koolen | Oliver Kopp | Axel Korthaus | Agnes Koschmider | Jan-Niklas Kramer | Oliver Krancher | Johann Kranz | Nane Kratzke | Yvonne Kreis | Barbara Krumay | Thomas Kude | Hannes Kuebel | Niklas Kühl | Michael Kunz | Jens F. Lachenmaier | Jens Lamprecht | Lesley Land | Michael Lang | Fabian Lang | Birger Lantow | Christoph Lattemann | Sven Laumer | Wool-Rim Lee | Hyejin Lee | Alexander Leischnig | Daniel Leonhardt | Henrik Leopold | Joerg Leukel | Christian Leyh | Jan-David Liebe | Kirsten Liere-Netheler | Aleck Lin | Sonia Lippe | Nikolaus Lipusch | Patrick Lohmann | Diana Lück | Thomas Ludwig | Andre Ludwig | Nele Lüker | Anh Luong | Heiner Lütjen | Antti Lyyra | Wolfgang Maaß | Tonja Machulla | Christian Maier | Markus Manhart | Jocha Märkle-hUß | Andrea Marshall | Gunnar Martin | Matthias Marx | Katharina Maßner | Christian Matt | Stefan Mau | Simon Mayer | Arne Meeue | Jan Meller | Patrick Mennig | Laura Menth | Johannes Merkert | Dirk Metzger | Mey Mark Meyer | Corvin Meyer-Blankart | Florian Michahelles | Andre Miede | Tobias Mirsch | Michael Möhring | Lars Mönch | Stefan Morana | Dominik Morar | Tina Morlok | Benedikt Morschheuser | Amir Mosavi | Florian Müller | Oliver Müller | Marcus Müller | Christoph Müller-Bloch | Jan Muntermann | Joschka Mütterlein | Thomas Myrach | Mario Nadj | Hari Suman Naik | Sivie Naimer | Arash Negahban | Gerard Neiditsch | Simon Nestler | Maria Neumann | Dirk Neumann | Juliane Neumann | Adelka Niels | Peter Niemeyer | Christina Niemöller | Volker Nissen | Verena Nitsch | Michael Nofer | Benedikt Notheisen | Thomas Nowey | Markus Nuettgens | Goran Oblakovic | Caroline Oehlhorn | Thuy Duong Oesterreich | Marek Opuszko | Uwe Ostermann | Sven Overhage | Sebastian Pape | Marcel Papert | Dominik Papies | Hannes Parbs | Jan Pawlowski | Norman Pelzl | Guenther Pernul | Christoph Peters | Mathias Petsch | Juergen Pfeffer | Alexander Pflaum | Bastian Pfleging | Christoph Pflügler | Melanie Pfoh | Gaurang Phadke | Vinh Pham | Henri Pirkkalainen | Ante Plazibat | Jens Pöppelbuß | João Porto de Albuquerque | Marco Pospiech | Tobias Potthoff | Michael Prilla | Wolfgang Prinz | Thorsten Proehl | Nicolas Pröllochs | Yu Pu | Carlos Rabes | Rene Ramacher | Matthias Rass | Madeleine Rauch | Ivan Razo-Zapata | Michael Rebstock | Beke Redlich | Victoria Reibenspiess | Kai Reimers | Jochen Reiner | Stefan Reinheimer | Diana Renner | Christopher Rentrop | Jonas Repschläger | Christian Reuter | Marcel Rhyn | Marc Riar | Sebastian Richter | Janek Richter | Rene Riedl | Bodo Rieger | Dennis Riehle | Andreas Riener | Uwe Riss | Raphael Rissler | Susanne Robra | Robert Rockmann | Erin Roehrer | Maximilian Röglinger | Katarzyna Rollert | Christoph Rosenkranz | Kristina Rosenthal | Richard Rößler | Angela Roth | Hannes Rothe | Dominik Rüegger | Julian Runge | Ben Ryder | Philipp Saalmann | Stefan Sackmann | Abdolkarim Sadrieh | Christina Sarigianni | Jacques Sauvé | Silvia Schacht | Thomas Schäffer | Siegfried Schallenmueller | Christian Scheiner | Raphael Schilling | Sebastian Schlauderer | Christian Schlereth | Hannes Schlieter | Alexander Schmid | Rainer Schmidt | Nikolaus Schmidt | Albrecht Schmidt | Lars Schmidt-THieme | Joachim Schnurbus | Bjoern Schoedwell | Andreas Schoknecht | Hendrik Scholta | Michael Scholz |

Johannes Schöning | Eric Schoop | Maximilian Schreieck | Guido Schryen | Petra Schubert | Reinhard Schuette | Peter M. Schuler | Carsten Schultz | Thomas Schulz | Thomas Schulze | René Schumann | Christoph Schütz | Florian Schwade | Andre Schweizer | Widura Schwittek | Christian Seel | Thomas Seeling | Christin Seifert | Martin Semmann | Thomas Setzer | Tobias Seyffarth | Katerina Shapoval | Giorgi Shuradze | Markus Siepermann | Mario Silic | Alexander Simons | Griselda Sinanaj | Oliver Skroch | Theresa Spickers | Patrick Spieth | Thorsten Spitta | Kai Spohrer | Benjamin Spottke | Thorsten Staake | Jeannette Stark | Philipp Staudt | Dennis Steininger | Lucas Stich | Steffen Stock | Jan Stockhinger | Julia Stoffregen | Susanne Strahringer | Christine Strauss | Stefan Strohmeier | Carola Stryja | Eric Sucky | Daniel Szopinski | Baseem Taher | Ioannis Tarnanas | Yiea-Funk Te | Timm Teubner | Frank Teuteberg | Stefan Thalmann | Malte Thiede | Oliver Thomas | Verena Tiefenbeck | Jan Tiegel | Kevin Tierney | Böhmann Tilo | Ingo J. Timm | Peter Tinschert | Julia Troll | Apostolos Tryfonidis | Sanja Tumbas | Tuure Tuunanen | Falk Uebernickel | Meike Ullrich | Ömer Uludag | Han van der Aa | Dirk van der Linder | Dirk Veiel | Vasilis Ververis | Arthur Vetter | Kristin Vogelsang | Ulrike von Baer | Aliona von der Trenck | Matthias von Entreß-Fürsteneck | Sergej von Janda | Denis Vuckovac | Michael Wächter | Thomas Wagenknecht | Benjamin Wagner vom Berg | Lena Waizenegger | Andy Weeger | Anne Weibert | Julian Weidinger | Matthias Weidlich | Michael Weiler | Peter Weimann | Christoph Weinert | Markus Weinmann | Anette Weisbecker | Markus Welle | Helena Wenninger | Stefan Wenzel | Sigrid Wenzel | Karl Werder | Christoph Weskamp | Lauri Wessel | Thomas Widjaja | Alexander Wieneke | Manuel Wiesche | Stephan Wilczek | Christoph Willing | Tristan Wimmer | Raphael Wimmer | Runhild Wineke | Axel Winkelmann | Maike Winkler | Matthias Winkler | Patrick Winter | Marco Wirth | Jakob Wirth | Matthias Wissotzki | Alexandre Witte | Rikka Wittstock | Krzysztof Wnuk | Matthias Wolf | Stefan Wolpert | Felix Wortmann | Pawel W. Wozniak | Volker Wulf | Jochen Wulf | Nancy Wünderlich | Emrah Yasasin | Iryna Yevseyeva | Christian Zagel | Eruani Zainuddin | Robert Zepic | Yuchen Zhao | Hans-Dieter Zimmermann | Alfred Zimmermann | Andreas Zolnowski | Christoph Zott | Patrick Zschech | Stefan Zwicklbauer

Best Paper Award

Nominierungen *(aufsteigend anhand der Paper-ID sortiert)*

Do We Really Want Blockchain-Based Accounting? Decentralized Consensus as Enabler of Management Override of Internal Control
Nadine Rückeshäuser

"I did use it!" - Assessing subjective vs objective cognitive artifact usage
Stefan Morana, Julia Kroenung, Alexander Maedche

What Teams Need to Be Clear about – an Activity Theoretical Perspective on Shared Understanding in Health IS Implementation
Andy Weeger, Annkatrin Ott-Schwenk

Best Paper Award Winner
Over-Paid Search: When Bricks-and-Mortar Retailers Should Not Use Paid Search
Darius Schlangenotto, Dennis Kundisch

Recombinant Service System Engineering
Daniel Beverungen, Hedda Lüttenberg, Verena Wolf

Ermittlung des Virtualisierungspotenzials von Beratungsleistungen im Consulting
Volker Nissen, Henry Seifert

Best Student Paper Award

Nominierungen *(aufsteigend anhand der Paper-ID sortiert)*

Award Winner des Liechtenstein Chapter of the Association for Information Systems
A Picture is Worth a Thousand Words: Visual Model Evaluation in Data Science Applications
Dennis Eilers, Michael H. Breitner

Examining the role of changing organizational culture on IT employees' commitment in M&A – Insights of a case study with a German software company
Aybala Ilgili

Best Student Paper Award Winner
Reconstructing the Giant: Automating the Categorization of Scientific Articles with Deep Learning Techniques
David Dann, Matthias Hauser, Jannis Hanke

The Core Capabilities of Green Business Process Management – A Literature Review
João Carlos Maciel

Design Prinzipien für Microlearning Crowdsourcing-Systeme. Konzept für audiovisuelle Mediengestaltung
Damian Kogga, Nicole Krawietz, Fatih Cevik, Sven Brandau, Mahei Manhai Li

Best Prototype & Experiment Award

Nominierungen *(aufsteigend anhand der Reihenfolge der Präsentation sortiert)**

Best Prototype & Experiment Award Winner
One Plug at a Time - Designing a Peer-to-Peer Sharing Service for Charging Electric Vehicles
Jan Hendrik Betzing, Moritz von Hoffen, Florian Plenter, Friedrich Chasin, Martin Matzner, Jörg Becker

Feldexperiment zur Wirksamkeit von konkretem vs. abstraktem Eco-Driving Feedback
André Dahlinger, Felix Wortmann, Verena Tiefenbeck, Ben Ryder, Bernahrd Gahr

Mobiles Lernen für China - eine iterative Prototypenentwicklung
Sissy-Josefina Ernst, Andreas Janson, Matthias Söllner ***First Runner-Up***

Architecture and Evaluation Design of a Prototypical Serious Game for Business Information Visualization
Christian Karl Grund, Michael Schelkle, Max Hurm

Towards a Software Prototype Supporting Automatic Recognition of Sketched Business Process Models
Manuel Zapp, Peter Fettke, Peter Loos ***Second Runner-Up***

Live Query - Visualized Process Analysis
Dennis M Riehle, Steffen Höhenberger, Rainer Cording, Patrick Delfmann

Empowering Smarter Fitting Rooms with RFID Data Analytics
Matthias Hauser, Matthias Griebel, Jannis Hanke, Frédéric Thiesse

Enhancing Asthma Control through IT: Design, Implementation and Planned Evaluation of the Mobile Asthma Companion
Peter Tinschert, Filipe Barata, Tobias Kowatsch

*Alle Beiträge des Tracks waren nominiert, da der Award im Rahmen der Prototype & Experiment Session durch das Publikum vergeben wurde.

Best Associate Editor Award

Nominierungen *(aufsteigend anhand des Nachnamens sortiert)*

Eva Bittner

Regine Dörbecker

Paul Drews

Martin Matzner

Gerhard Satzger (Best Associate Editor Award Winner)

Thomas Setzer

Till Winkler

Best Reviewer Award

Nominierungen *(aufsteigend anhand des Nachnamens sortiert)*

Dominik Bilgeri

Ricardo Büttner

Paul Drews (Best Reviewer Award Winner)

Phillip Haake

Henrik Leopold

Christina Niemöller

Jens Pöppelbuß

Maximilian Röglinger

Inhaltsverzeichnis

Track 13: Prototype & Experiment

Autonomie im Internet der Dinge: Anforderungen an die Gestaltung autonomer Agenten in Cyber-physischen Systemen

Christian Janiesch[1] und Valentin Nentwich[2]

[1] Julius-Maximilians-Universität Würzburg, Würzburg, Germany
christian.janiesch@uni-wuerzburg.de
[2] Audi Planung GmbH, Neckarsulm, Germany
valentin.nentwich@audi-planung.de

Abstract. Durch die Digitalisierung in Gesellschaft und Industrie findet eine wachsende Autonomisierung bzw. Automatisierung der Wertschöpfungssysteme über Sensoren und Effektoren, sogenannte Cyber-physische Systeme, statt. Diese Automatisierung erlaubt eine schnellere und häufig auch fehlerfreiere Arbeit in vernetzten Systemen, bspw. im Bereich Industrie 4.0. Es ist derzeit allerdings weder möglich noch sinnvoll jedweder Maschine volle Autonomie zu gewähren, so dass es nötig ist, genau zu definieren, in welchem Rahmen und mit welchen Zielen eine Maschine selbstbestimmt handeln darf. Allerdings gibt es für die Gestaltung von autonomen Maschinen im CPS derzeit keinen allgemein anerkannten Ansatz oder Methodik, wie solch ein Autonomie-Konzept umgesetzt werden könnte. Basierend auf einer Literaturanalyse der IT-Autonomie-Forschung werden 12 Anforderungen für die Umsetzung von autonomen Agenten in CPS erarbeitet. Es werden die Eigenschaften eines Agenten sowie dessen Interaktion mit dem Menschen über Autonomiestufen detailliert aufgearbeitet. Das resultierende Meta-Modell kann als Basis für die Gestaltung von Agenten, bspw. über eine Modellierungssprache, dienen.

Keywords: *Internet der Dinge, Cyber-physische Systeme, IT-Autonomie, Modellierung, Agent*

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Janiesch, C.; Nentwich, V. (2017): Autonomie im Internet der Dinge: Anforderungen an die Gestaltung autonomer Agenten in Cyber-physischen Systemen, in: Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1-15

1 Einleitung

Digitalisierung ist ein Leittrend in Gesellschaft und Industrie, der eine Kernherausforderung für die Wirtschaftsinformatik darstellt. In der Produktion wird in diesem Zusammenhang von Industrie 4.0 als dem Organisationsprinzip der Zukunft gesprochen. Allgemeiner spricht man von selbststeuernden Objekten, die Internet-ähnliche, vernetzte Strukturen zur Kommunikation nutzen, dem sogenannten Internet der Dinge [1]. So soll letztendlich eine wachsende Autonomisierung bzw. auch Automatisierung der Wertschöpfungssysteme über Sensoren und Effektoren, in sogenannten Cyber-physischen Systemen (CPS), stattfinden.

Das starke Wachstum an Sensoren und deren Vernetzung fördert dabei die Erzeugung und Verbreitung von Daten und wird folglich aufgrund der Menge und der Geschwindigkeit, in der Daten heute entstehen, zu einer großen Herausforderung für Entscheidungsträger. Diese sind letztendlich durch die Vielzahl an Informationen und die Beschränktheit ihrer kognitiven Fähigkeiten nicht in der Lage, adäquate und zeitnahe Entscheidungen für alle durchgeführten Analysen zu treffen, ohne dabei Optimierungspotentiale zu verschenken und ohne Fehler zu machen [2].

Ein möglicher Lösungsansatz besteht darin, den Grad der Autonomie von Maschinen zu erhöhen. Mit anderen Worten ausgedrückt bedeutet dies, eine Maschine zu befähigen Entscheidungen als selbstständiger Agent zu treffen und dadurch die Kontrolle über Situationen und Handlungen zu übernehmen. Allerdings gibt es für die Gestaltung von autonomen Maschinen im CPS derzeit keinen allgemein anerkannten Ansatz oder keine Methodik, wie ein Autonomie-Konzept umgesetzt werden könnte.

Der Schwerpunkt dieses Artikels liegt auf der Systematisierung von Forschungsansätzen der IT-Autonomie. Basierend auf einer Literaturanalyse wird ein Konzept für die Umsetzung von Autonomiestufen und autonomen Agenten in CPS erarbeitet und die Basis für deren Spezifikation, bspw. über eine grafische Modellierungssprache, vorbereitet, welche die Modellierung von Autonomie ermöglichen würde. Das bereitgestellte Konzept schließt die Lücke zwischen der abstrakten Idee von autonomen CPS und der praktischen Gestaltung solcher Systeme in der Realität.

Im Folgenden werden zunächst Grundlagen zum Internet der Dinge und CPS vorgestellt. Darauffolgend werden die Kernaspekte hinsichtlich IT-Autonomie aufgezeigt. Diese allgemeingültigen Erkenntnisse bilden das Fundament der abschließend vorgestellten Grundlagen für eine Spezifikation des autonomen Verhaltens von Agenten. Sie gliedern sich zunächst in mehrere Anforderungen und werden anschließend in einem Meta-Modell umgesetzt und diskutiert. Das Papier schließt mit einem Ausblick auf weitere notwendige Arbeiten zur Gestaltung der Autonomie in CPS.

2 Grundlagen

2.1 Cyber-physische Systeme und Internet der Dinge

In der Literatur finden sich zahlreiche Erläuterungen des Begriffs Cyber-physisches System, eine anerkannte Definition liefern diese jedoch nicht. [3, 4] prägten zuerst den

Begriff des Internet der Dinge für Computer, die nicht nur eigenständig Informationen sammeln, sondern dazu befähigt werden, aus den gesammelten Informationen zu lernen und die Welt durch ihr stetig kumuliertes Wissen selbst wahrzunehmen und über ein Netzwerk, bspw. das Internet, kommunizieren zu können.

[5] differenziert darauf aufbauend CPS als physische, biologische und/ oder bautechnische Systeme, die über eine Recheneinheit integriert, überwacht oder gesteuert werden. Die Recheneinheit ist dabei ein verteiltes, eingebettetes System. [6] hebt zudem die Interdependenzen von Recheneinheiten und physischen Prozessen in Feedbackschleifen hervor. [7] ergänzen die Notwendigkeit einer oder mehrerer multi-modaler Mensch-Maschine-Schnittstellen zur Kommunikation und Steuerung von CPS und die weltweite Nutzbarkeit von Daten und Diensten für lokale Operationen.

Zusammenfassend bezeichnen CPS also die Kopplung von informations- und softwaretechnischen Komponenten mit mechanischen bzw. elektronischen Komponenten, die über eine Kommunikationsinfrastruktur, wie zum Beispiel das Internet, in Echtzeit miteinander kommunizieren. Diese Komponenten können in unterschiedlichem Grade automatisiert und autonom handeln, woraus Konsequenzen für die Mensch-Maschine-Interaktion entstehen [2].

2.2 (IT-)Autonomie

Der Begriff Autonomie besitzt im Allgemeinen die zwei Bedeutungen Unabhängigkeit und Selbstständigkeit. Selbstständigkeit beschreibt die Fähigkeit einer Entität, bzw. eines Agenten, für sich selbst zu sorgen, während Unabhängigkeit für die Freiheit einer Entität vor der Kontrolle einer von außen einwirkenden Instanz steht [8].

Im Kontext der IT kann Autonomie als eine nichtfunktionale Systemeigenschaft gesehen werden, die sich auf einzelne Funktionen übertragen lässt. Autonomie wird hierbei als Spezialisierung der Adaptivität betrachtet, der Fähigkeit zu verstehen, eine Zielstellung in Bezug auf eine Eingabemenge zu erfüllen und dabei die eigenen Strukturen an die Änderungen in der Umwelt anzupassen. Es wird eine Bindung von Autonomie an eine Funktion erreicht, indem die autonome Funktionserfüllung als Basiseigenschaft aufgefasst wird, mehrere Systeme miteinander in Bezug zu setzen [9].

Autonomie steht in einem engen Zusammenhang mit Automatisierung, welche eine maschinelle Handlung ohne menschliche Unterstützung darstellt. Es ist demnach ein Mittel für die Ablösung von Aufgaben eines Menschen [10]. Durch Automatisierung kann eine höhere Geschwindigkeit sowie eine bessere Effizienz bei der Aufgabenbewältigung erreicht werden. Weiterhin kann der physische und geistige Aufwand für einen Menschen reduziert werden.

IT-Autonomie kann den Menschen insbesondere bei Koordinationsaufgaben entlasten und so die Geschwindigkeit oder Sicherheit erhöhen. Sie lässt sich nicht nur gänzlich auf einzelne funktionale Aspekte eines Agenten übertragen, sondern durch eine graduelle Abstufung unterschiedlicher Ebenen der Interaktion zwischen Menschen und Maschinen kennzeichnen [11]. Diese Abstufung kann als Kontinuum zwischen der vollständigen Autonomie, bei der Maschinen Entscheidungen selbstständig gemäß ihrer eigenen Zielstellungen treffen, und dem größten kognitiven sowie physischen, menschlichen Einsatz auf der anderen Seite verstanden werden [2].

Wir verstehen unter Autonomie keine reine Systemeigenschaft, welche sich anhand des Informationsgradienten und des emergenten Verhaltens feststellen lässt [9]. Vielmehr rücken Aspekte der Interaktion zwischen einem Menschen und einer Maschine sowie der dadurch festgelegte Grad vorhandener Autonomie in einem System in den Vordergrund (siehe Kapitel 3.3). Abgedeckt wird hierbei das Spektrum vollständiger Kontrolle der Zielerreichung durch sowohl den Menschen als auch der Maschinen sowie das Kontinuum an Möglichkeiten, welches zwischen beiden Extremen liegt.

3 Autonomie in Cyber-physischen Systemen

3.1 Eigenschaften autonomer Agenten

Agenten sind eine Entität, wie bspw. Maschinen aber auch der Mensch, die ihre Umwelt über Sensoren wahrnehmen und mittels Effektoren auf diese Umwelt einwirken können [12]. Autonome Agenten handeln dabei entsprechend ihrer eigenen Agenda [13]. Die wichtigsten Bestandteile eines Agenten sind folglich Sensoren für die Sinneswahrnehmung und Effektoren für das Ausführen von Aktionen auf eine Umwelt. Eine Umwelt bezeichnet dabei den Wirkungskreis eines Agenten. Nach dem Verständnis von [14] stellt eine Umwelt ein System dar, welches aus einer Vielzahl von Agenten besteht. Hierbei kann ein System selbst wiederum ein Agent sein, welcher ein Teil eines größeren Systems respektive Agenten ist.

Beispiele für Sensoren menschlicher Agenten sind deren Augen und Ohren. Als ihre Effektoren werden häufig Hände und Mund genannt. Für künstliche Agenten zählen bspw. Kameras, Temperaturfühler oder digitale Schnittstellen zu Sensoren. Effektoren können Teile der Robotik aber, auch akustische Signale oder elektronisch übertragene Nachrichten sein.

Die Implementierung der Logik, Zielstellung und Wertvorstellung in Bezug auf die Informationsverarbeitung eines Agenten lassen [12] offen. Sie unterscheiden für die konkrete Definition genannter Eigenschaften folgende vier Typen von Agenten:

- **Agent mit Reflexen:** Agiert entsprechend einer vorgegebenen Menge von Regeln. Befindet sich die wahrgenommene Welt in einem definierten Zustand, wird die dem Zustand zugeordnete Aktion ausgeführt.
- **Agent mit internen Zuständen:** Erweitert einen einfachen Agenten mit Reflexen, so dass nicht nur der aktuelle Umweltzustand, sondern auch die Vergangenheit und die verfolgte Entwicklung des Umweltzustandes relevant für die auszuwählende Aktion des Agenten sind.
- **Agent mit Zielen:** Erweitert einen Agenten mit internen Zuständen um die Fähigkeit anhand von Zielstellungen den Einfluss von Aktionen auf seine Umwelt abzuwägen. Eine Zielvorgabe beschreibt den Zustand der Umwelt, der durch einen Agenten erreicht werden soll. Hierbei wird nicht genauer spezifiziert, wie der Zielzustand erreicht werden kann [15].
- **Agent mit Nutzenfunktion:** Erweitert einen Agenten mit Zielen um die Fähigkeit, den Nutzen für seine möglichen Handlungsalternativen festzustellen und die entsprechend der Nutzenfunktion beste Alternative zu wählen.

Durch das Hinzufügen des Aspektes der Autonomie erhalten Agenten die Fähigkeit, selbstständig zu entscheiden, wann sie welche Aufgaben wie ausführen und wann sie mit wem, auf welche Art und Weise, zu diesem Zweck kommunizieren. Hierdurch sind Agenten fähig, auf vorhergesehene und unvorhergesehene Ereignisse aus ihrer Umwelt zu reagieren, erwartete und unerwartete Probleme zu lösen und Teil einer Organisation mit dynamischen Strukturen zu sein [15, 16]. [13] geben eine Übersicht von Eigenschaften, welche für die Klassifizierung von Agenten (siehe Tabelle 1) herangezogen werden kann. Ihrer Auffassung zufolge erfüllen autonome Agenten zumindest die ersten vier Eigenschaften dieser Übersicht: Reaktivität, Autonomie, Ziel-Orientierung und Kontinuität. Reaktivität in Form einer zeitnahen Reaktion auf asynchrone, externe Reize ist eine Schlüsseleigenschaft eines autonomen Agenten. Deshalb sollte die Informationsverarbeitung in Echtzeit erfolgen [17].

Tabelle 1. Eigenschaften zur Klassifizierung von Agenten (in Anlehnung an [13])

Eigenschaft	*Beschreibung*
Reaktiv	Reagiert zeitnahe auf Veränderungen in der Umwelt
Autonom	Übt Kontrolle über seine eigenen Aktivitäten aus
Ziel-orientiert	Keine einfache Reaktion auf die Umwelt
Temporär	Kontinuierlich laufender Prozess
Kommunikativ	Kommuniziert mit anderen Agenten, eventuell Menschen eingeschlossen
Lernend	Verändert sein Verhalten basierend auf Erfahrungen
Mobil	Fähigkeit, sich selbstständig auf eine andere Maschine zu transportieren
Flexibel	Aktionen folgen keinem Skript
Charakter	Glaubhafte „Persönlichkeit“ und emotionaler Zustand

[18] heben ergänzend die sozialen Fähigkeiten insbesondere die Kommunikation mit anderen künstlichen oder menschlichen Agenten, als zentrale Eigenschaft hervor. [19] untersuchen soziale Interessen von Agenten in Bezug auf die Entscheidungsfindung. Sie unterscheiden zwischen folgenden drei Formen sozialer Kompetenzen:

- **Eigennutz:** Ein Agent stellt die eigene Zielstellung respektive seine eigenen Vorteile über den globalen Nutzen der Gruppe.
- **Hilfsbereitschaft:** Ein Agent wählt Alternativen, welche aus Sicht der Gruppe vorteilhaft sind. Hierbei sind auch Alternativen möglich, die keinen individuellen Vorteil für den einzelnen Agenten bringen.
- **Kooperation:** Agenten wählen Alternativen, welche einen positiven Nutzen für die Gruppe kooperierender Agenten bringen. Zwar ist eine Wahl von Alternativen möglich, die einen individuellen Nachteil des Agenten nach sich ziehen, allerdings wird dieser Nachteil von einem Vorteil übertroffen, der durch eine in einem Team ausgeführte Aktion erzielt wird.

Vorangehende Ausführungen bezüglich der Konzepte zu Autonomie und zu autonomen Agenten zeigen, dass Informationsverarbeitung auf der einen Seite und Interaktivität auf der anderen Seite ausschlaggebend für die Bestimmung des Autonomielevels von Maschinen sind. Vergleicht man die Menge der innerhalb der Grenzen eines Systems verarbeiteten Informationen mit der über die Systemgrenzen hinweg ausgetauschten Informationsmenge, lässt sich der Informationsgradient als Indikator des Autonomiegrades eines Systems bestimmen [9].

Bezüglich der Informationsverarbeitung stützen sich [20, 21] auf die vier Funktionen Überwachung, Erzeugung, Auswahl und Implementierung. Diese können wie folgt auf den Prozess der Informationsverarbeitung übertragen werden: Systemstatus wahrnehmen, mögliche Reaktionen ableiten, Aktionen auswählen und Aktionen umsetzen. Während [20] den Grad der Autonomie über die Zuordnung der Grundfunktionen zu Menschen und Maschinen festlegen, beschreiben [21] eine Divergenz des Autonomiegrades innerhalb der einzelnen Funktionen selbst.

[22] liefern eine Taxonomie acht möglicher Kommunikationsmuster zwischen Menschen und Maschinen, im Speziellen Robotern.

- **1 → 1:** Der Mensch nimmt eine steuernde Rolle ein, während die Maschine einerseits Informationen aus der Umwelt wahrnimmt und an den Menschen weiterleitet und andererseits die durch den Menschen vorgegebenen Aktionen ausführt.
- **1 → 1+1+…:** Zwei oder mehr Maschinen werden von einem Menschen gesteuert. Der Mensch ist für die Aufteilung der Aufgabe und die Koordination der Ausführung zuständig. Die Maschinen arbeiten unabhängig voneinander ihre Teilaufgaben ab.
- **1 → n:** Zwei oder mehr Maschinen werden von einem Menschen gesteuert. Die Maschinen nehmen die Anweisung durch den Menschen als Gruppe wahr und koordinieren die Unterteilung der Aufgabe sowie deren Ausführung selbst.
- **n → 1:** Zwei oder mehr Menschen erarbeiten in einem Team Aktionen für eine Maschine und übermitteln diese nach gemeinsamer Abstimmung an eine Maschine zur Ausführung.
- **n → 1+1+…:** Zwei oder mehr Menschen erarbeiten in einem Team Aktionen für zwei oder mehr Maschinen. Die Menschen sind für die Aufteilung der Aufgabe und die Koordination der Ausführung zuständig. Die Maschinen arbeiten dann unabhängig voneinander ihre Teilaufgaben ab.
- **n → n:** Zwei oder mehr Menschen erarbeiten in einem Team Aktionen für zwei oder mehr Maschinen und übermitteln diese nach gemeinsamer Abstimmung an die Maschinen zur Ausführung. Die Maschinen nehmen die Anweisung durch den Menschen als Gruppe wahr und koordinieren die Unterteilung der Aufgabe sowie deren Ausführung selbst.
- **1+1+… → 1:** Zwei oder mehr Menschen übermitteln unabhängig voneinander Aktionen an eine Maschine. Die Maschine übernimmt die Koordination eingehender Aufgaben.
- **1+1+… → n:** Zwei oder mehr Menschen übermitteln unabhängig voneinander Aktionen an eine Gruppe von Maschinen. Die Maschinen entscheiden innerhalb ihrer Gruppe über die Zuweisung der Aufgaben an einzelne Gruppenmitglieder sowie über die Prioritäten bei der Aufgabenbearbeitung.

Die Fähigkeit der Gruppenbildung weist Parallelen zu komplexen adaptiven Systemen auf [23], in denen Agenten miteinander interagieren, aber auch lernen und sich anpassen. Eine weitere sich daraus ergebende Eigenschaft ist die Selbstähnlichkeit.

Zur Übertragung von Informationen ist ein Kommunikationsmedium erforderlich. [24] beschreibt Varianten der Kommunikation eines Menschen mit einer Maschine und vice versa. Diese orientieren sich an den klassischen fünf Sinnen Sehen, Hören, Riechen, Fühlen und Schmecken.

[25] bezeichnen Vertrauen als Basis der Verlässlichkeit bei der Kommunikation. Vertrauen ist immer dann ein notwendiger Faktor, wenn eine Bewertung von Handlungsalternativen aufgrund von Unsicherheiten und hoher Komplexität nicht möglich ist. Der Grad des Vertrauens dient dann als ein Indikator dafür, ob eine Entscheidung akzeptiert oder selbst getroffen wird. Attribute für die Vertrauenswürdigkeit von Kommunikation sind bspw. Aufmerksamkeit, Verlässlichkeit, Konformität und Sicherheit.

3.2 Anforderungen an autonome Agenten

Agenten agieren in einer Umwelt. Sie müssen Reize mittels Sensoren wahrnehmen, Aktionen wählen und diese Aktionen durch den Einsatz eigener Effektoren an ihre Umwelt zurückgeben. Die Umwelt kann dabei als holonisches Multi-Agenten-System betrachtet werden. Das bedeutet, dass ein Agent aus der Vogelperspektive beobachtet wie ein einzelnes Teil mit sichtbaren Sensoren und Effektoren wirkt. Im Querschnitt dieses Agenten wird deutlich, dass sich der Aufbau des Hauptagenten aus einer Vielzahl von Subagenten zusammensetzt. Es besteht eine rekursive Beziehung zwischen Agenten, die eine beliebige Granularität zulässt.

Anforderung 1: Die Spezifikation muss hierarchische Anordnungen von autonomen Agenten abbilden.

Ein Agent besitzt beliebig viele Sensoren und beliebig viele Effektoren. Mit Effektoren gibt ein Agent Informationen eines bestimmten Typs ab. Der Agent selbst oder andere Agenten in seiner Reichweite nehmen diese Informationen mit den jeweils entsprechenden Sensoren wahr.

Anforderung 2: Die Spezifikation muss abbilden, welche Arten von Sensoren und Effektoren von Agenten unterstützt werden und welche Sensoren die Datenobjekte welcher Effektoren wahrnehmen können.

Die Richtung einer Interaktion bestimmt den Datenfluss zwischen den Teilnehmern an einer Interaktion. Hierdurch wird auch die Abhängigkeit zwischen Agenten sowie die Gesamtstruktur eines Netzwerkes aus Agenten ausgedrückt.

Anforderung 3: Die Spezifikation muss die Richtungsdefinition einer Interaktion abbilden, so dass deren kommunikatives Verhalten sichtbar gemacht wird.

Ein Agent ist genau dann autonom, wenn dieser Aufgaben selbstständig ausführen kann. Hierzu zählt das Treffen relevanter Entscheidungen bzgl. Ressourcen und Ausführungszeitpunkten.

Anforderung 4: Die Spezifikation muss einfache und komplexe Regeln, welche das Handeln eines Agenten beschreiben, abbilden.

Ergänzend müssen die Entscheidungen auf Basis des internen und externen Zustandes getroffen werden.

Anforderung 5: Die Spezifikation muss verschiedene Zielstellungen, Nutzenfunktionen und Zustände abbilden.

Die Anforderung, in Echtzeit auf Stimuli zu reagieren, zeigt insbesondere bei Menschen, dass die Auswertung von Sensordaten nur in einem beschränkten Maße möglich ist. Die als Aufmerksamkeitskapazität bezeichnete Fähigkeit beschreibt sowohl eine Beschränkung des Kurzzeitgedächtnisses als auch der Phasen Wahrnehmung, Entscheidungs- und Aktionsauswahl sowie Aktionsdurchführung. Aus ökonomischen oder technologischen Gründen kann eine vergleichbare Beschränkung auch auf Maschinen zutreffen.

Anforderung 6: Die Spezifikation muss eine kapazitive Beschränkung von Funktionalitäten durch ein Konstrukt ähnlich der Aufmerksamkeit abbilden.

Unter dem Aspekt der Gruppierung respektive der Teambildung lassen sich mehrere Agenten zu einer Gruppe zusammenfassen. Beispielsweise übergibt ein Mensch eine Aufgabe an eine Gruppe von Maschinen, welche die Arbeitsaufteilung und Aufgabenbewältigung selbst koordiniert. Wie beschrieben können Agenten auch aus beliebig vielen Subagenten aufgebaut und durch Zielstellungen, Nutzenfunktionen, Regeln und Zustände spezifiziert werden. Diese müssen nicht konfliktfrei sein.

Anforderung 7: Die Spezifikation muss den Aspekt der sog. sozialen Rationalität abbilden und Formen sozialer Kompetenzen darstellen sowie eine Möglichkeit abbilden, Agenten unterschiedlichen Typs zu Gruppen zusammenzufassen. Das gilt auch für Agenten desselben Typs, wenn deren Kommunikationsverhalten eindeutig bestimmt ist. Dies erfordert die Bildung von Agentenklassen, bei denen die Anzahl möglicher, konkreter Instanzen individuell festgelegt werden kann.

Sicherheit ist einerseits ein Mechanismus, der Korruption bei der Nachrichtenübertragung verhindert und somit bewusste Manipulationen und Angriffe von Dritten auf Datenströme erschwert. Andererseits wird dem Sicherheitsaspekt die Eigenschaft der Zuverlässigkeit bei der Nachrichtenübermittlung zugeordnet.

Anforderung 8: Die Spezifikation muss Vertrauens-Parameter einer Beziehung respektive eines Datenübertragungskanals, wie Aufmerksamkeit, Verlässlichkeit, Konformität und Sicherheit, abbilden.

Die Beispiele zu den allgemeinen Koordinationsprotokollen sowie die Ansätze aus dem Bereich des Spezialfalls der Supervisory Control [26] zeigen, dass strukturelle Vorgaben für die Koordination von Agenten über deren Zustände und Regeln abgebildet werden können. Folglich entstehen aus dem Bereich der Koordination keine gesonderten Anforderungen.

3.3 Anforderungen an die Interaktion autonomer Agenten

Es gibt mehrere Konzepte die sich mit der Einteilung von Autonomie in unterschiedliche Stufen bei der Interaktion zwischen Mensch und Maschine beschäftigen [2, 20, 21, 26-30]. Tabelle 2 fasst die verschiedenen Konzepte zusammen und bildet die entsprechenden Stufen ab. Eine detaillierte Besprechung der verwandten Arbeiten findet sich in Kapitel 5. Eine eindeutige stufenartige Staffelung, welche sich zwischen den beiden Extremen der vollständigen Autonomie einer Maschine S (Stufe 1) oder eines Menschen M (Stufe 19) bewegt, ist nicht möglich.

Die beiden Extremfälle auf Stufe 1 und 19 beschreiben jeweils Situationen, in denen entweder ein Mensch oder eine Maschine die vollständige Kontrolle über die Entscheidungsfindung und die Aktionsdurchführung besitzen. Die bereits für autonome Agenten identifizierten Anforderungen genügen, um beide Sachverhalte adäquat abzubilden. Folglich entstehen aus den Randbereichen keine gesonderten Anforderungen.

Tabelle 2. Übersicht der Autonomiestufen

Nr	*Beschreibung*	*Quelle*
1	S trifft Entscheidung komplett und führt Aktionen autonom aus, keine Funktion durch M	[2, 20, 21, 26, 27, 30]
2	S trifft Entscheidung komplett und führt Aktionen autonom aus, M überwacht S und greift ggf. steuernd ein	[20, 27, 28]
3	S trifft Entscheidung komplett und führt Aktionen autonom aus, M überwacht bestimmte Bereiche von S und greift ggf. steuernd ein	[29]
4	M trifft Entscheidung und führt Aktion autonom aus bis zur Übergabe an S, dann führt S Aktionen autonom aus	[20, 26]
5	M trifft Entscheidung und übergibt Aktion zur Ausführung an S	[20]
6	S trifft Entscheidung und führt Aktion solange aus bis M ein Veto einlegt	[2]
7	S wählt Aktion und gibt M beschränkte Zeit für Veto	[21, 26, 30]
8	S wählt Aktion und führt nur auf Zustimmung von M aus	[20, 21, 26, 27, 30]
9	S bietet M alle Alternativen zur Auswahl an	[21]
10	S bietet M vorausgewählte Alternativen zur Auswahl an	[2, 21, 26, 27, 30]
11	S bietet M genau eine Alternative an, M kann dennoch ablehnen	[2, 21, 26, 27, 30]
12	S und M schlagen Alternativen vor, M wählt beste Alternative aus	[20]
13	M schlägt S Alternativen vor, S wählt die beste Alternative aus und führt diese aus	[20]
14	S führt Aktion aus und benachrichtigt M nach eigenem Ermessen	[21, 26]
15	S führt Aktion aus und benachrichtigt M auf Nachfrage von M	[21, 26, 30]
16	S führt Aktion aus und benachrichtigt M immer	[21, 30]
17	M trifft Entscheidungen und steuert S	[28]
18	M trifft Entscheidungen und steuert, aber S übernimmt Teile der Steuerung und führt diese autonom aus	[28]
19	M trifft Entscheidung komplett und führt Aktionen autonom aus, keine Funktion durch S	[20, 21, 27, 30]

Stufe 2 und 3 beschreiben eine Situation, in der einer der beiden Interaktionspartner Aktionen auswählt und ausführt, während der andere Partner die Rolle einer überwachenden Instanz einnimmt und in kritischen Situationen steuernden Einfluss auf die Ausführung nimmt. Die beiden Stufen unterscheiden sich im Grad der Überwachung, welcher sich über die Menge der überwachten Funktionen festlegen lässt. Beide Stufen entsprechen dem Konzept der Supervisory Control und den Anforderungen, die mit diesem Konzept verknüpft sind (siehe oben).

Eine weitere Gruppe bilden die 4. und 5. Stufe. Bei beiden liegt die Autorität einer Entscheidung bei einem Menschen. Beide Stufen unterschieden sich durch den Grad der Aufteilung einer Aufgabe bei deren Ausführung. Durch das Konzept der autonomen Agenten wurde bereits eine Anforderung definiert, welche Datenobjekte als Teil des hier auszuarbeitenden Meta-Modells beschreibt, die zwischen zwei oder mehreren Agenten ausgetauscht werden. Der Fokus dieser beiden Gruppen liegt auf dem Austausch von Aufgaben und Teilaufgaben.

Anforderung 9: Die Spezifikation muss Aufgabenlisten als speziellen Typ eines Datenobjektes abbilden.

Ein ähnlicher Fall ergibt sich aus den Stufen 9 bis 11, welche unterschiedliche Mengen an Alternativen, die einem Agenten von einem anderen Agenten vorgeschlagen werden, beschreiben: (1) alle Alternativen, (2) eine Auswahl an Alternativen sowie (3) genau eine Alternative. Stufe 12 und 13 beschreiben das gleiche Verhalten aus gegensätzlicher Perspektive.

Anforderung 10: Die Spezifikation muss Datenobjekte des Typs Aufgabenlisten entsprechend der beschriebenen Abstufung abbilden.

Stufe 6 bis 8 beschreiben die Situation, in der eine Maschine selbstständig Aktionen auswählen und ausführen kann. Die Stufen unterscheiden sich darin, dass ein Mensch unterschiedliche Möglichkeiten besitzt, durch sein Veto die Ausführung einer Aktion zu unterbrechen oder gar zu verhindern.

Anforderung 11: Die Spezifikation muss Datenobjekte zur Zustimmung und Ablehnung von Aufgaben abbilden.

Eine weitere Kategorie kann aus den Stufen 14 bis 16 abgeleitet werden. Diese beschreiben die Situation, dass ein Agent eine Aktion ausführt und dabei einen weiteren Agenten auf unterschiedliche Art über seinen aktuellen Zustand benachrichtigt: (1) nach eigenem Ermessen benachrichtigen, (2) immer benachrichtigen und (3) auf Anfrage benachrichtigen.

Anforderung 12: Die Spezifikation muss den Typ der Benachrichtigung abbilden.

Die verbleibenden Stufen 17 und 18 zeigen starke Ähnlichkeit zu den Stufen 2 und 3. Der Unterschied zwischen beiden Gruppen liegt in der aktiven beziehungsweise in der passiven Haltung des menschlichen Agenten. Auf der einen Seite nimmt dieser eine passive, beobachtende Rolle ein und wird nur in kritischen Situationen steuernd aktiv. Auf der anderen Seite ist der menschliche Agent durchgehend aktiv, trifft zu einem gewissen Grad Entscheidungen und steuert den zweiten Agenten. Auf abstrakter Ebene betrachtet, macht die Frequenz der Eingriffe den Unterschied zwischen den Stufen aus, welche sich über die Komplexität der zu steuernden Aufgaben ergibt. Folglich kann dieser Sachverhalt über das Formulieren von Regeln für die beteiligten Agenten abgebildet werden und stellt keine weiteren Anforderungen.

4 Meta-Modell zur Spezifikation autonomer Agenten

Der folgende Abschnitt definiert die abstrakten Bestandteile einer Sprache, ein Meta-Modell, das die Spezifikation von Autonomie erlauben soll. Abbildung 1 stellt den Aufbau eines Agenten in drei Kategorien dar: Bestandteile, kennzeichnende Elemente und beschreibende Elemente. Diese modellierten Informationen könnten dann für die Charakterisierung der Autonomie in CPS, bspw. von (Sub-)Prozessen im Bereich Business Process Management oder von Event Processing Agents bzw. Networks im Bereich

Complex Event Processing, verwendet werden, um zu signalisieren, wie autonom sich diese Prozesse bzw. Analyse-Netzwerke in Bezug auf ihre Umwelt verhalten sollen.

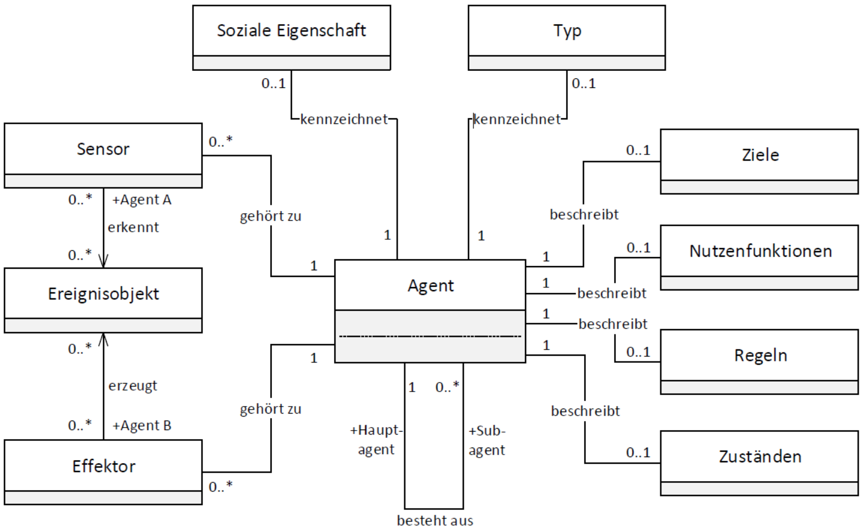


Abbildung 1. Übersicht Meta-Modell

Die Kategorie Bestandteile umfasst die Sensoren und Effektoren eines Agenten. Im traditionellen Sinne ist ein Sensor immer genau einem Agenten zugeordnet. Maschinen sind jedoch grundsätzlich in der Lage gemeinsam auf Sensoren und Effektoren zuzugreifen. Weiterhin kann ein Agent beliebig viele Sensoren zugeordnet bekommen. In der gleichen Beziehung steht ein Agent zu seinen Effektoren. In beiden Fällen sind auch die Extreme möglich, dass einem Agenten keine Sensoren oder keine Effektoren zugeordnet werden. Im Meta-Modell ist dies implizit über die Hierarchisierung von Agenten abgebildet.

Es gibt grundsätzlich sechs Kanäle über die Effektoren und Sensoren kommunizieren: visuell, auditiv, olfaktorisch, taktil, gustatorisch und generisch. Die Einteilung der ersten fünf Kanäle orientiert sich hierbei an den von Schmauk als klassische Modalitäten eingeordneten fünf Sinnen. Der generische Sensor dient als offener Typ, der über zusätzliche Parameter näher bestimmt werden kann und somit alle diejenigen Kanäle abbilden kann, welche nicht durch eine der anderen fünf Klassen spezifiziert werden können. Hierzu zählen bspw. elektronische Kommunikation von Maschinen oder Hirnströme bei Menschen.

Zwei oder mehrere Agenten können in einer Beziehung stehen, wenn sie Sensoren und Effektoren besitzen, die kompatibel zueinander sind, d. h. wenn Ereignisobjekte erzeugt werden, die von einem der Sensoren des anderen Agenten erkannt werden können. Eine Beziehung zwischen zwei Agenten kann durch optionale Parameter beschrieben werden. Diese beschreibenden Parameter können von Agenten bei deren Entscheidungsfindung herangezogen werden und können somit Einfluss auf die Entscheidung selbst haben. Ein Beziehungsparameter offenbart grundlegende Eigenschaften, die das

Verhältnis zwischen zwei Agenten betreffen. Beziehungsparameter spezialisieren sich zu Aufmerksamkeit, Verlässlichkeit, Konformität und Sicherheit.

Die Aufmerksamkeit bezeichnet wie oben eingeführt eine kapazitive Obergrenze an Informationen, die in den verschiedenen Stufen der Verarbeitung wahrgenommen werden kann. Wir nutzen dieses Konzept, um den von einem Effektor erzeugten Anteil der Ereignisobjekte einer Beziehung zu beschreiben, der von einem Sensor erkannt wird, oder um technische Limitationen bei der Auswertung zum Ausdruck zu bringen. Verlässlichkeit und Konformität leiten sich aus den Konzepten bezüglich des Vertrauens bei der Interaktion zwischen Mensch und Maschine ab. Der Unterschied zwischen den beiden Parametern liegt in ihrer Wirkungsrichtung. Verlässlichkeit beschreibt das Vertrauen eines wahrnehmenden Agenten gegenüber dem sendenden Agenten. Es ist also ein Indikator dafür, zu welchem Grad die Signale zur Entscheidungsfindung von einem wahrnehmenden Agenten herangezogen werden. Die Konformität wirkt entgegengesetzt zu der Verlässlichkeit und beschreibt somit das Vertrauen eines sendenden Agenten gegenüber dem empfangenden Agenten. Sicherheit beschreibt als letzter Aspekt die Vertrauenswürdigkeit des Kommunikationskanals.

Die zweite Kategorie enthält die kennzeichnenden Elemente eines Agenten. Hierzu zählen einerseits die sozialen Eigenschaften und andererseits die unterschiedlichen Typenausprägungen eines Agenten. Eine gesonderte Betrachtung der sozialen Eigenschaften zu den unterschiedlichen Agententypen ist notwendig, um soziale Eigenschaften als Indikator für das Auflösen von Konflikten zwischen den Zielen, Nutzenfunktionen, Regeln und Zuständen eines Hauptagenten zu seinen Subagenten bzw. Konflikten zwischen Agenten in einem Netzwerk zu nutzen und somit das spezifische, autonome Verhalten eines Agenten zu zeigen. Die Kennzeichnung durch den Typ eines Agenten gibt Aufschluss über das allgemeine Verhalten bei der Ereignisverarbeitung. Einem Agenten kann maximal eine soziale Eigenschaft zugeordnet werden. Die Assoziation zwischen einem Agenten und seinem Typ folgt der gleichen Kardinalität wie bei den sozialen Eigenschaften.

Eine dritte Kategorie bilden die vier beschreibenden Elemente Ziele, Nutzenfunktionen, Regeln und Zustände. Die Summe der Elemente respektive deren Zusammenspiel beschreiben das autonome Verhalten eines Agenten. Die Kardinalitäten der Assoziation eines Agenten mit den jeweiligen Elementen dieser Kategorie entspricht den bereits beschriebenen Kardinalitäten der kennzeichnenden Merkmale. Folglich sind auch diese Elemente für die Spezifikation optional und zeigen primär an, ob ein Agent durch eine Menge von Zielen, Nutzenfunktionen, Regeln und Zuständen beschrieben wird oder nicht.

Eine weitere Eigenschaft eines Agenten, die durch Abbildung 1 dargestellt ist, ist eine Assoziation zu sich selbst. Diese reflexive Assoziation zeigt an, dass sich ein Hauptagent in beliebig viele Subagenten zerlegen lässt.

5 Verwandte Arbeiten

[26] stellen eine zehnstufige Skala für die Unterscheidung verschiedener Grade der Automatisierung bei der Entscheidungsfindung vor. Zwischen den beiden Extremen der

vollständigen Entscheidungskompetenz liegen Stufen, welche durch das Bilden von Interaktionsmustern zwischen Mensch und Maschine differenziert werden.

Auch [21] beschreiben eine graduelle Abstufung von Autonomie in zehn Stufen. Jede dieser Stufen ist hierbei auf die einzelnen Phasen der menschlichen Informationsverarbeitung anwendbar. Neben der Stufe der vollständigen Autonomie werden die Stufen nach den drei Kategorien Art der Information, Vetorechte und Anzahl der angebotenen Handlungsalternativen differenziert. [30] reduzieren diese Skala auf acht Stufen. Die Autoren gehen hierbei speziell auf den Fall einer durch den Computer ausgewählten Handlungsalternative ein.

[31] beschreibt ein Modell, welches die grundlegenden Eigenschaften aus dem Konzept der Supervisory Control abbildet. Er beschreibt eine fünfstufige Skala für die Bestimmung des Automatisierungsgrades in Bezug auf den Prozess der Informationsverarbeitung. [29] identifiziert sieben Arten der Beeinflussung gegenseitiger Autonomie durch Veränderung von Systemparametern. Bisherige Ansätze werden dahingehend erweitert, dass autonomes Verhalten nicht nur von der eigenen Autorität über das Treffen von Entscheidungen abhängt, sondern auch von weiteren Entitäten im System beeinflusst und geprägt wird.

[32] schlägt 13 Eigenschaften eines Agenten vor, anhand derer die autonome Abhängigkeit zu seiner Umwelt festgestellt werden kann. [14] bauen auf diesen Eigenschaften der Dimensionierung von Autonomie auf und zeigen in ihren Untersuchungen Autonomie in Bezug auf Multi-Agenten-Organisationen.

[27] zeigt vier unterschiedliche Grade von Autonomie auf. [33] greifen diese Skala auf und vergeben für die autonome Abstufung die folgenden Bezeichnungen: keine Systemunterstützung, Entscheidungsunterstützung, konsensuelle und überwachte, künstliche Intelligenz sowie vollständige Autonomie. [20] erweitern die Skala zu einer Taxonomie mit zehn Stufen. Diese reichen von manueller Kontrolle bis zu vollständiger Autonomie.

[28] klassifiziert unterschiedliche Grade von Autonomie: Überwachung und logisches Denken, Planung des Fahrwegs, Kollisionsverhinderung und einfache Steuerung. Auch diese Abstufung reicht von der vollständigen Kontrolle des Menschen bis zur Überwachung von Zielvorgaben durch den Menschen.

[19] beschreiben in ihrem Ansatz das Verhalten von Agenten in Gruppen. Sie knüpfen Autonomie an die soziale Bereitschaft eines Agenten, welche über die Zielstellung eines Agenten an die Maximierung eines individuellen oder gesellschaftlichen Nutzens gebunden ist. [8] wählen ebenfalls einen Ansatz, welcher Autonomie unterschiedliche Dimensionen zuordnet. Sie beschreiben die Anpassung von Autonomie entlang gewählter Dimensionen durch das Einschränken oder das Erweitern von Berechtigungen, Fähigkeiten, Möglichkeiten und Verpflichtungen eines Partners, welcher sich in einer Beziehung zur Erfüllung gemeinsamer Aufgaben befindet.

[2] untersuchen wie sich die automatisierungsbedingte Leistungsfähigkeit von Menschen bei unterschiedlichen Automatisierungsstufen verändert. Die Autoren weisen auf einen Kosten-Nutzen-Kompromiss hin, der sich daraus ergibt, dass zwar Routine-Aufgaben bei korrekter Arbeitsweise auch bei veränderlicher Arbeitsbelastung schneller erledigt werden können, aber die Performance und das Situationsbewusstsein deutlich sinkt, wenn das Entscheidungsunterstützungssystem fehlerhaft arbeitet.

6 Zusammenfassung und Ausblick

Aufbauend auf einer umfangreichen Literaturstudie zur IT-Autonomie wurden 12 Anforderungen für die Spezifikation autonomer Agenten in CPS-Systemen bestimmt. Diese wurden übersichtlich in einem abstrakten Meta-Modell zusammengefasst, welches die Möglichkeit eröffnet, die Autonomie von Agenten eindeutig zu beschreiben. Es ist zunächst bewusst abstrakt gehalten und kann für verschiedene Komponten eines Mensch-Maschine-Systems verwendet werden, wie bspw. ein CPS-System, das Menschen sowie Maschinen mit oder ohne physischer Präsenz (bspw. virtualisierte Prozess- oder Analyse-Systeme) enthält.

Die nächsten Schritte werden sich mit der Umsetzung des abstrakten Modells in eine konkrete Notation und die Verknüpfung derer mit existierenden Modellierungssprachen beschäftigen. Außerdem werden wir Entwurfsmuster für die Mensch-Maschinen-Kommunikation basierend auf den vorgestellten Autonomie-Leveln entwickeln, um die Implikationen für den Kommunikationsablauf besser darstellen zu können.

Literaturverzeichnis

1. Kopetz, H.: Real-Time Systems: Design Principles for Distributed Embedded Applications. 2nd edn. Springer, New York, NY (2011)
2. Onnasch, L., Wickens, C., Li, H., Manzey, D.: Human Performance Consequences of Stages and Levels of Automation: An Integrated Meta-Analysis. Journal of the Human Factors and Ergonomics Society 56, 476-488 (2014)
3. Ashton, K.: That “Internet of Things” Thing. RFID Journal (2009)
4. Weiser, M.: The Computer for the 21st Century. Scientific America 265, 94-104 (1991)
5. Gill, H.: From Vision to Reality: Cyber-Physical Systems. In: Proceedings of the HCSS National Workshop on New Research Directions for High Confidence Transportation CPS: Automotive, Aviation, and Rail, Tyson’s Corner, VA (2008)
6. Lee, E.A.: Cyber Physical Systems: Design Challenges. Technical Report No. UCB/EECS-2008-8 University of California, Berkeley, CA (2008)
7. Geisberger, E., Broy, M.: agendaCPS: Integrierte Forschungsagenda Cyber-Physical Systems. Springer, Heidelberg (2012)
8. Bradshaw, J., Feltovich, P., Jung, H., Kulkarni, S., Uszok, W.T.A.: Dimensions of Adjustable Autonomy and Mixed-initiative Interaction. Agents and Computational Autonomy. LNCS, vol. 2969, pp. 17-39. (2004)
9. Richling, J., Werner, M., Jaeger, M., Mühl, G., Heiß, H.-U.: Autonomie in verteilten IT-Architekturen. De Gruyter, München (2011)
10. Chestnut., H.: Automation: What it is and What are the Problems it Poses. Automatica 1, 241-252 (1963)
11. Parasuraman, R., Riley, V.: Humans and Automation: Use, Misuse, Disuse, Abuse. Human Factors: The Journal of the Human Factors and Ergonomics Society 39, 230-253 (1997)
12. Russell, S., Norvig, P.: Articial Intelligence: A Modern Approach. Pearson, Harlow (2014)
13. Franklin, S., Graesser, A.: Is it an Agent, or Just a Program?: A Taxonomy for Autonomous Agents. Intelligent Agents III: Agent Theories, Architectures and Languages. LNCS vol. 1193, pp. 21-35. Springer, Berlin (1997)

14. Schillo, M., Fischer, K.: A Taxonomy of Autonomy in Multiagent Organisation. Agents and Computational Autonomy. LNCS vol. 2969, pp. 68-82. Springer, New York, NY (2004)
15. Munroe, S., Luck, M.: Agent Autonomy Through the 3M Motivational Taxonomy. Agents and Computational Autonomy. LNCS vol. 2969, pp. 55-67. Springer, New York, NY (2004)
16. Jennings, N.: Agent-based Computing: Promise and Perils. In: Proceedings of the 16th International Joint Conference on Articial Intelligence, pp. 1429-1436. Stockholm (1999)
17. Brustoloni, J.: Autonomous Agents: Characterization and Requirements. Technical Report, Pittsburgh, PA, Carnegie Mellon University (1991)
18. Wooldridge, M., Jennings, N.: Intelligent Agents: Theory and Practice. The Knowledge Engineering Review 10, 115-152 (1995)
19. Kalenka, S., Jennings, N.: Socially Responsible Decision Making by Autonomous Agents. In: Proceedings of the 5th International Colloquium on Cognitive Science, pp. 135-149. Springer, Donostia-San Sebastián (1999)
20. Endsley, M., Kaber, D.: Level of Automation Effects on Performance, Situation Awareness and Workload in a Dynamic Control Task. Ergonomics 42, 162-192 (1999)
21. Parasuraman, R., Sheridan, T., Wickens, C.: A Model for Types and Levels of Human Interaction with Automation. IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans 30, 286-297 (2000)
22. Yanco, H., Drury, J.: Classifying Human-robot Interaction: An Updated Taxonomy. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, pp. 2841-2846. The Hague (2004)
23. Holland, J.H.: Studying Complex Adaptive Systems. Journal of Systems Science and Complexity 19, 1-8 (2006)
24. Agah, A.: Human Interactions with Intelligent Systems: Research Taxonomy. Computers & Electrical Engineering 27, 71-107 (2000)
25. Lee, J., See, K.: Trust in Automation: Designing for Appropriate Reliance. Human Factors: The Journal of the Human Factors and Ergonomics Society 46, 50-80 (2004)
26. Sheridan, T., Verplank, W.: Human and Computer Control of Undersea Teleoperators. Technical Report Massachusetts Institute of Technology, Cambridge, MA (1978)
27. Endsley, M.: The Application of Human Factors to the Development of Expert Systems for Advanced Cockpits. In: Proceedings of the Human Factors and Ergonomics Society 31st Annual Meeting, pp. 1388-1392. New York, NY (1987)
28. Jipp, M.: Levels of Automation: Effects of Individual Differences on Wheelchair Control Performance and User Acceptance. Theoretical Issues in Ergonomics Science 15, 479-504 (2014)
29. Sheridan, T.: Adaptive Automation, Level of Automation, Allocation Authority, Supervisory Control, and Adaptive Control: Distinctions and Modes of Adaptation. IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans 41, 662-667 (2011)
30. Sheridan, T., Parasuraman, R.: Human-automation Interaction. Reviews of Human Factors and Ergonomics 1, 89-129 (2005)
31. Sheridan, T.: Supervisory Control. In: Salvendy, G. (ed.) Handbook of Human Factors, pp. 1025-1052. Wiley, New York, NY (1997)
32. Castelfranchi, C.: Founding Agent's 'Autonomy' On Dependence Theory. In: Proceedings of the 14th European Conference on Articial Intelligence, pp. 353-357. Amsterdam (2000)
33. Endsley, M., Kiris, E.: The Out-of-the-loop Performance Problem and Level of Control in Automation. Human Factors: The Journal of the Human Factors and Ergonomics Society 37, 381-394 (1995)

Do We Really Want Blockchain-Based Accounting? Decentralized Consensus as Enabler of Management Override of Internal Controls

Nadine Rückeshäuser

Institute of Computer Science and Social Studies, Department of Telematics, Freiburg, Germany
nadine.rueckeshaeuser@iig.uni-freiburg.de

Abstract. Research proposing the application of blockchain technology in accounting assumes the utilization of decentralized consensus mechanisms based on the exertion of scarce resources (Proof-of-Work; PoW), leading to the validation of transactions without the need of any third party. Together with the blockchain, a shared database, PoW is expected to lead to nearly immutable and, therefore, fraud-resistant, real-time financial registers. This conclusion must be reconsidered, taking into account recurrent top-management involvement in accounting scandals, often conducted through deliberate exposures of internal and external control systems. This paper asserts that blockchain-based accounting using PoW-based consensus paves the way for the suspension of controls by the management, since exerting the majority of computer power is easier than circumventing internal and external control systems in conventional accounting systems. Alternatives to PoW must be considered for blockchain-based accounting that prevents the management from conducting fraud and, thereby, qualifies the blockchain for its application in accounting.

Keywords: *Decentralized Consensus, Blockchain Customization, Blockchain-Based Accounting, Accounting, Management Override of Controls*

1 The Blockchain: The New Cure All?

Today, blockchain technology in combination with decentralized consensus mechanisms (DCMs) and its application in various business sectors is on everyone's lips. Blockchains are shared databases that are maintained and verified amongst actors that participate in a network, ensuring digital transparency and confidence of records of information without a trusted third party [26]. Whereas the financial sector was an early adopter, the demand for the technology has increased over the past years and comes from diverse industries, such as health care or logistics. This demand is not surprising, given the blockchain's ability to enable decentralized autonomous business models, defined by self-governed programs through decentralized governance and collective consensus [27]. In particular, this allows the execution of

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Rückeshäuser, N. (2017): Do We Really Want Blockchain-Based Accounting? Decentralized Consensus as Enabler of Management Override of Internal Controls, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 16-30

Turing-complete codes for so-called smart-contracts, leading to self-executing programs that automatically enforce properties of digital contracts [35].

One major business sector expected to benefit from the features of the blockchain is accounting. In particular, blockchains may facilitate the maintenance of permanent and timely records of financial transactions [36]. Its decentralized and transparent nature further implies potential immutability, meaning that financial records cannot be altered ex post and, if so, the probability of detection will be very high [2]. Thus, blockchain-based accounting could possibly rule out the conduction and concealment of improper accounting methods, illicit structuring of transactions and financial database manipulations [16]. The possibility of blockchain-based accounting, therefore, has recently become an intensively discussed issue, not only in an industrial [2, 34] but also in an academic context [8, 25, 36]. The growing interest in this topic is also reflected by the formation of several start-ups offering blockchain-based services for decentralized bookkeeping, such as *Factom* [37] or *Scorechain* [38]. Overall, the application of blockchain technology in the context of accounting could be conducive to the industry, which is still mainly based on standardized technology such as computer assisted audit techniques [2]. As the digitization of accounting is still in its infancy, the application of blockchain technology may leads to the technological progress needed.

However, industrial and academic advocates of blockchain-based accounting seem to neglect the still present and well-known challenges of proper accounting that is top-management involvement in accounting fraud [17, 32]. The severity of this topic gets obvious when looking, for instance, at fraud incidents in the United States, where accounting fraud conducted by the management amounts to 89 percent of all financial statement fraud cases of public companies [3]. The following paper investigates whether the blockchain is qualified for an application in accounting. To this end, it is investigated how the management is able to conduct fraud and whether the proposed intra-corporate blockchain application is able to decrease the opportunities to commit fraud. Accordingly, it is assumed that there exist incentives for the management to commit fraud, however, there might be technical mechanisms for its actual prevention. Most of all, this implies an investigation of DCMs regarding their ability to impede the management from conducting accounting fraud.

The paper is structured as follows: The next section provides a case study of the Comroad accounting scandal and investigates the used manipulation and concealment techniques. Based on this case study, a generalization of the fraud pattern and the relationships between internal and external control systems is deducted, using additional scientific literature to support the identified relationship in the case of Comroad. Section 3 subsequently presents a layer model for blockchain customization, on which basis a scenario for blockchain-based accounting systems is developed. Using the scenario as well as the general fraud pattern and opportunities to commit fraud, identified in section 1, various DCMs are investigated concerning their ability to serve as technical mechanism for fraud prevention.

2 How to Conduct Accounting Fraud: A Case Study & Analysis

Accounting fraud is the deliberatively attempt to prepare and disseminate material that misstates a company's financial situation [32]. Involvement of the top management, such as Chief Executive Officers (CEOs) and/or Chief Financial Officers (CFOs), in accounting fraud (hereafter: management accounting fraud, MAF) is frequently observed [3]. Thereby, MAF includes either direct involvement of top-management in conducting accounting fraud or indirect involvement by convincing or enforcing the provision of fraud by other parties. To identify the requirements on blockchains in respect to MAF prevention, a case study of the Comroad accounting fraud scandal and a generalization of this case for a further analysis are provided.

2.1 The Comroad Accounting Scandal

Comroad was a German telematics service provider, who developed worldwide applicable, server-based traffic systems. These systems were sold to trading partners, whereas retailers offered the systems as well as complementary services to end-costumers [15]. Comroad entered the international trading floor in the beginning of 1999, whereas it sales quadrupled at the end of this year, compared to its prior year's level of DM4.6 million. Afterwards, the company exhibited exorbitant growth perspectives, despite overall negative trends in the industry. In particular, Comroad forecasted an increase of sales to DM250 million in 2002 [15].

Comroad's success story, however, turned out to be one of the major accounting scandals of publicly traded firms in Germany. Sales developments were the result of numerous fictitious transactions, for which Comroad invented commercial relationships with non-existing trading partners, amounting to €19.9 million as declared in Comroad's financial annual report [11]. One of those trading partners was a company named *VT Electronics*, which was allegedly in charge of the production and deliver of board computers on behalf of Comroad. However, *VT Electronics* was only collecting money from likewise fictitious end-customers. For the purpose of concealment, payment from end-customers was cleared with the production costs of equipment of *VT Electronics* and by down payment for further hardware and possible retained surpluses. The only task of Comroad was to prepare invoices and to pretend that invoices were send to end-customers. Comroad stated additional transactions with various other Asian trading partners following a similar fraud pattern [15]. Surprisingly, the illicit practices of Comroad were not detected over a period of three years and despite of various controls in accordance to national and international legal requirements (e.g. the Germany Stock Corporation Act (AktG), Euro-Bilanzgesetz (EuroBilG)) as well as standards for accounting (e.g. IFRS). Thus, in the following the manipulation and concealment techniques of Comroad will be discussed.

2.2 The Manipulation and Concealment Techniques of Comroad

Dorin [15] describes several incidents of MAF and also the previously presented case study of Comroad by using the so-called *swiss cheese* model. The model shows, how systems may breakdown due to human intended or unintended as well as technical failures [28]. Accordingly, MAF may be conducted despite the existence of several legally required and/or voluntarily implemented firewalls, i.e. internal and external control systems as well as technical precautions. Perpetrators of fraud are able to circumvent those controls and security measures by using the system's deficiencies (loopholes) for their own benefit [15].

According to German regulations, the publicly traded company Comroad was managed by a two-tier board structure consisting of the management, responsible for the oversight of day-to-day business operations, and the board of directors, responsible for oversight of the management and acting as final authority with respect to decision making [1]. Despite this top-down approach of control, additional internal controls are legally required, for example, according to the AktG [13], the Act for Control and Transparency in the Corporate Sector (KonTraG) [7], as well as auditing standards such as IDW PS 261 [18]. Internal controls are measures and methods adapted to safeguard assets of a company as well as to check the accuracy of bookkeeping [7]. However, there exist no specified requirements for the corporate-specific design of internal control systems in the German legislation. In general, the board of directors is under legal obligation to monitor the implementation and development of an adequate internal control system, which may includes internal auditors and/or an audit committee [7].

In the case of Comroad, it seems obvious that neither the board of directors nor the internal control system was sufficient to prevent the deduction of accounting fraud. Particularly, the CEO of Comroad was able to bypass and to suspend the internal control system – a practice called management override of internal controls [9] – by staffing the board of directors with his wife, who was involved in the fraudulent activities, thereby undermining the board's independence. As a consequence, the board tolerated the illicit practices [15] as well as the internal auditor, who received monetary remuneration for the maintained silence [10]. Despite the arrangements within the company, the establishment of a close relationship to the external auditor KPMG, by which both parties received mutual advantages, offset external controls. Lastly, financial statement users such as investors, the stock market as well as supervisory authorities, were misled and deceived, as they relied to a great extend on the audited and certified financial statements attested by KPMG [11, 15].

2.3 Generalization: Dependencies in Control Systems and Accounting Fraud

The described hierarchy of the company, the external and internal control system as well as the associated relationships between the control systems and organs in the context of the Comroad scandal are transferred to an arbitrarily chosen stock company i, to which the German Stock Corporation Act applies. The identified loopholes in the control systems of Comroad are generalized and crosschecked by the economic literature, among others [1, 3, 6, 15] as well as [32], and combined in Figure 1. For example, existing theoretical or empirical work about the management's influence on the board of director's independence [1] and other relationships [4, 9] were reviewed, by which the management is able to exert control over the board. Solid lines indicate that there is a broad consensus about the indicated relationship in the examined literature, whereas dashed lines emphasize controversies. The influence of the management on the board of directors is determined by a variety of factors, such as the geographically disparity or career perspectives of board members. However, the strong influence of the management on the board of directors in the case of Comroad can be substantiated by a large part of the literature and is, therefore, assumed within Figure 1. It should be noted that the emphasized relationships are not generally true, but may be part of the problem when considering the emergence of MAF.

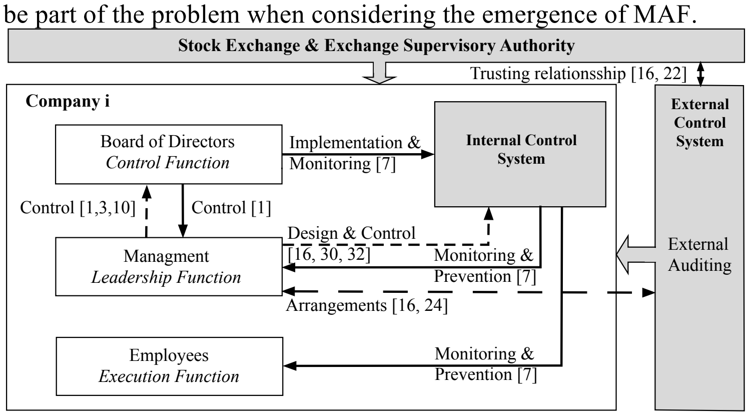


Figure 1. Stylized illustration of a company's control system and dependencies

Analyzing Figure 1, the existence of circular references between the management, the board of directors and the internal control system get apparent. In particular, if the board of directors depends partly or completely on the management and if the internal control system is determined by the management, than there exist no internal mechanism that might prevent the management from conducting accounting fraud by the exertion of effective controls. This observation is supported by the findings of Sawyer [30] as well as Caplan [9], noticing that the management will always be able to override internal controls, especially because they are able to choose the strength of these systems through its influence on the board. Moreover, if management override of controls happens, there is no obvious reason for external auditors to revise their evaluation of management integrity [9, 22], leading to additional negative effects for the effectiveness of controls. Despite the effects of external auditors are controversial discussed, the observations of [9, 22] coincide with those in the case of Comroad, where the assurance service of external auditor seems to have deteriorated and, therefore, was not able to deter management fraud [15]. Moreover, these inefficiencies are expected to exert further negative effects on the external control system, that is, first of all, external auditors, as well as the market, which typically trust (at least to a great extend) third-party financial audit [22].

Overall, this leads to the conclusion that the core of the management's ability to conduct and conceal accounting fraud is an inefficient internal control system resulting from a dependent board of directors, which is strongly influenced by the management. Based on this observation, a first step towards the prevention of MAF seems to be the strengthening of the independence of both mechanisms. This can be done in terms of impeding the management's influence on the internal control system and the board of directors as well as by the avoidance of interdependencies between those entities and control systems. Second, MAF can also be prevented by decreasing the probability of successful concealment of fraud through the covering up of tracks, i.e through database manipulations and the circumvention of technical precautions. Consequently, an effective strategy for the prevention of MAF must take into account an organizational as well as a technical perspective.

2.4 Could Blockchains Prevent Accounting Fraud?

In view of the above considerations, the suitability of the blockchain for accounting must be discussed, as several academic and industrial research papers propose this, e.g. [2, 8, 36]. Certainly, the blockchain in combination with decentralized consensus induces organizational transformation through the decentralization of single business processes and by the potential increased involvement of employees because of high transparency. For instance, decentralized consensus could potentially raise employee involvement in accounting issues and the validation of business transactions, leading to more diversified controls through the transparency induced by the blockchain. Financial transparency is a major issue in accounting and for the internal control system, which is concerned about the openness and availability of information [20] that could potentially be moderated by the blockchain. Moreover, facilitating the involvement of employees could solve a frequently mentioned problem in internal control systems, which are claimed to be design, using an excessive agency view that promotes a strong adversarial relationship between the management and shareholders, but leaves out the relationship of the management and employees [32]. Summarized, given the potential organizational changes induced be the blockchain, it can be concluded that it is worthwhile to have a closer look on the technology and the impact of organizational restructuring. From a technical view, the blockchain is expected to introduce immutability of data stored on the blockchain, a feature that is frequently mentioned not only in the context of its possible application in accounting, e.g. [2, 8, 25]. This argument is based on the assumption to apply a proof-of-work (PoW) based DCM, which is a cryptographic puzzle, consisting of solving a mathematical problem by the exertion of computer power. In particular, PoW is a mechanism to ration resource access in client-server relationships and consist in finding a byte string combined with a block header, which results in a cryptographic hash that can only be done by the exertion of computer power [12]. Given this assumption, the suitability of an application of the blockchain for accounting will be analyzed not only with respect to the organizational transformations but also concerning the applied DCM.

3 Blockchain Customization and Organizational Restructuring

In this section, the structure and possible customization of the blockchain will be discussed in the context of a business environment. Based on the blockchain design decisions, a scenario for blockchain-based accounting will be presented.

3.1 The Blockchain in a Business Environment: Structure and Customization

Using a very basic definition, a blockchain is a synchronized global log of events between nodes in a peer-to-peer network. Particularly, a blockchain is replicated at every node and assists nodes in reaching consensus on the state of all accounts [26]. Blockchains can be customized for special use cases and adjusted to business environments, which is illustrated by a layer model presented in Figure 3. It is assumed that the layers overlap and are partially interconnected. The layer model provides an overview on how blockchains may fit in and support a company by providing deployment choices and by enabling flexibility.

At the lowest layer the blockchain provides a **digital infrastructure**, called distributed ledger. This basic infrastructure consists of three elements: The peer-to-peer network consists of homogenous nodes and is characterized by the ability to exist without a central node, responsible for network control. In the context of the blockchain, each node keeps a complete replica of all data needed to independently verify the validity of any data that should be incorporated into the distributed ledger. Before data are incorporated they must be broadcasted through the network. After broadcasting, a common order over data has to be agreed among the nodes, which is a non-trivial problem in a distributed network, known as the Byzantines generals problem [21]. This problem is solved by a cryptographically puzzle and the exertion of computer power by which a particular target value must be found (PoW; note that this mechanism will be explained in greater detail later). After reaching consensus, data are logged and permanently stored in the distributed ledger. At this level the customization of the blockchain for a business environment and/or application may happens by the choice of general rules according to which consensus is found.

The second layer is characterized by the choice of different **deployment modes** that depend on the desired openness of the peer-to-peer network and the type of data validation. Blockchains can either feature permission-less access or permissioned access. In the former case, everyone is able to participate as node and no prior authorization is needed [25]. Contrarily, blockchains that are characterized by permissioned access pre-select their participating nodes, e.g. through white- or blacklisting and some type of gatekeeping mechanism [31]. Irrespective of the access type, the validation of the blockchain can be either performed in a decentralized way or by one particular or several nodes, i.e. centralized validation. Despite this sounds counterintuitive using a DCMs, centralization may stems from the fact that the validation is transmitted to a set of changing nodes (e.g. delegated proof-of-work) that are responsible for the validation, for instance, to avoid too much overhead and to allow for low latency. Contrarily, decentralized validation is characterized by the fact that all nodes in the network are able to validate data that should be incorporated into the blockchain. Typically, permission-less decentralized ledgers are featured by decentralized validation (e.g. Bitcoin) whereas permissioned ledgers use centralized validation. However, every other combination or hybrid form are conceivable [31]. This definition excludes unintended centralization, for instance, through the undesired accumulation of the majority of computer power in the case of PoW.

On the third level, the blockchain is shaped by system design decisions in regard to the **foundation and integration** of the envisaged application. This includes that rules within the application must be designed in accordance to the particular business process or compliance requirements. In contrast to the general rules mentioned in the first layer that, these *specific* rules state additional, technical feasible requirements. For example, these rules may obey requirements referring to a specific section of the AktG. Moreover, a service-oriented architecture (SOA) might be taken into accounting, if a blockchain application must be integrated within an existing enterprise systems [33]. However, most likely not only the interaction of the blockchain and decentralized consensus with other information systems must be considered, but also user interaction. Nevertheless these system design decisions are only exemplary and customization may include numerous other aspects.

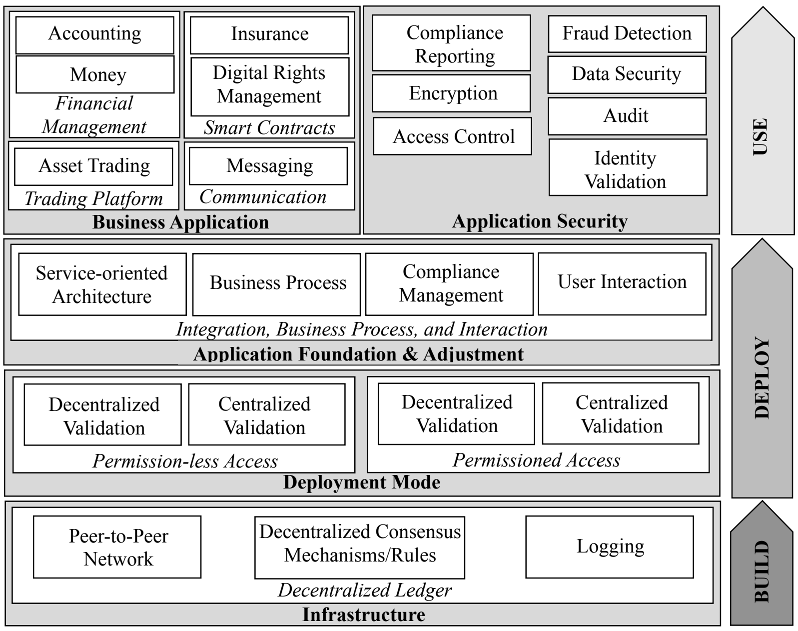


Figure 2. Blockchain layer model and customization in a business environment

Lastly, on the top layer particular **applications** are built on the basis of the preceding layer decisions. For example, smart contracts can be implemented for insurance services or digital rights management. However, as depicted in Figure 3, applications of the blockchain can relate to various business sectors, whereas this list is not exhaustive. Lastly, apart from the fact that the blockchain itself offers particular **security features** through cryptography, additional security mechanisms might be implemented on the upper layer. Depending on the concrete application, these mechanisms may range from additional data securing mechanism and fraud detection to audit, whereas, again, this list is not exhaustive.

3.2 Scenario: Blockchain-Based Accounting and Organizational Changes

Proposed applications of the blockchain for accounting vary significantly from joint registers [2] to intra-firm blockchain-based record keeping [8, 34, 36]. However, industrial and academic literature lack a description on the concrete implementation of the blockchain as well as application scenarios, on which basis the blockchain and the proposed DCM can be evaluated. Contrarily, this paper develops a scenario for a blockchain-based accounting system using the layer model for customization.

As depicted in Figure 3, the basic infrastructure of the proposed accounting system is the distributed ledger, where business transactions are referenced on as monetary value and not as tokens. The deployment model of the blockchain is a private blockchain maintained by a network of individuals within the company that validate transactions, here called intra-corporate blockchain. Particularly, intra-corporate blockchains are chosen in this paper, since full transparency of sensitive financial data to particular companies or - in an extreme scenario - to the general public could lead to severe losses in competitive advantages for an individual company. For example, lawfully discretionary accounting practices would be no longer feasible, which could be exploited by competitors, whose financial data are not completely transparent. Thus, the following scenario is inspired by the facts from the Comroad study and the obligation of German stock companies for publishing annual accounts (AktG), which not implies full and real-time transparency. A comparison of intra- and inter-corporate ledgers is purposely excluded by referring to the associated strong focus of this study.

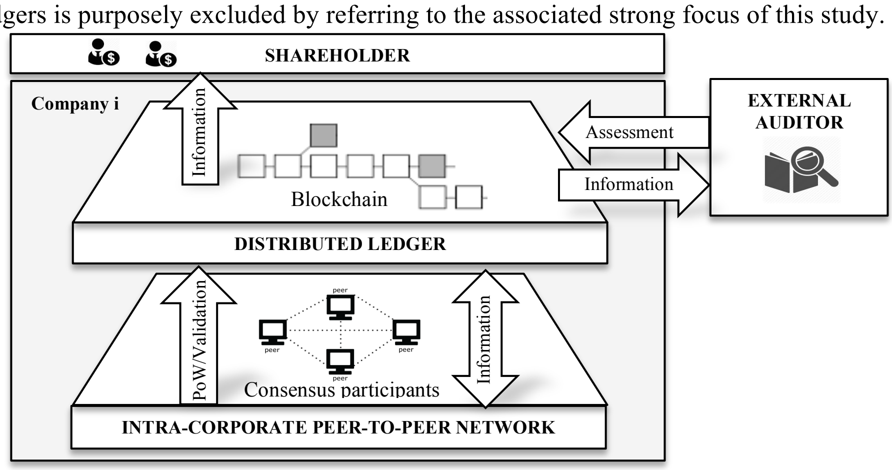


Figure 3. Scenario of a blockchain-based accounting

The network is assumed to consist of employees, especially, the accounting department, the management, and associated control entities that are the board of directors and an optional internal auditor or an audit committee, which together build the pool of consensus participants. Employees are likely to be enforced to participate on the consensus, as part of their work assignment. In contrast, executives and shareholders are assumed to act in their own interest and participate either because they want to influence the consensus protocol in a negative way, e.g. to conduct fraud, or in a positive way, as shareholder are likely to be concerned about the accuracy of the financial situation. Consensus is found in accordance to the PoW mechanism, which is assumed as DCM in all identified papers that propose an application of blockchain technology for accounting, e.g. [2, 8, 36]. In this system, consensus will only be found if transactions are in accordance to the pre-specified rules. Consensus participants will reject transactions that are not compliant. Valid transactions are subsequently logged and serve as publicly available source of information within the company and to particular outsiders (e.g. external auditors). Simultaneously, consensus participants are the source of information by conducting transactions via the accounting system and broadcasting it to the rest of the network for validation.

4 Can We Prevent Management Accounting Fraud?

In the following, PoW as well as alternative DCMs will be investigated and assessed concerning their ability to prevent MAF. According to the previously presented scenario the decentralized consensus cannot be separated from the peer-to-peer network. Thus, it is acknowledged that there exist threats that result from the peer-to-peer network. Related attacker scenarios are, among others, Sybil attacks, Eclipse attacks, Byzantines Joint attacks as well as Churn attacks [14]. Secure blockchain-based accounting system must account for those attacks. However, given numerous works dealing with the security of peer-to-peer networks, it is assumed that there exist mechanisms to provide a considerable security level for the network. Thus, in the following the focus lies on DCMs and their ability to prevent MAF.

4.1 Management Override of Controls and Proof-of-Work Based Consensus

PoW is a mechanism to rationing resource access in client-server relationships, consisting of solving a mathematical puzzle, using computer power [5]. Particularly, PoW consists in finding a byte string, called nonce that combined with the block header, results in a cryptographic hash with a given number of leading zero bits. A block contains all transactions, which have been committed on the previous block. Finding a nonce, can only be done by calculating the hash of the block for all possible nonces [12]. In addition, each block references to the preceding block, which hash has to be known, meaning that blockchains represent consensus over the history of data stored on the blockchain. The history is considered as true, when it deploys the longest chains, conforming to the exertion of the most power exertion. Thus, if someone wants to revert the history, an alternative reality must be created (blockchain fork), which occurs if not all nodes agree on the same blockchain header [12]. The blockchain fork will only be accepted if it becomes longer than the already existing blockchain, which implies the exertion of a huge amount of computer power, starting from the point that should to be altered (51% attack). This requires not only computational power but also faster data processing than the rest of the network [26].

While PoW seems to provide a reasonable level of security in large networks, small-scale networks have been proofed to remain vulnerable of 51% attacks [5]. In particular, this holds for intra-corporate blockchain-based accounting systems, where the management is potentially able to deliberately reach the majority of computer power. Without convincing, enforcing or circumventing exiting internal and external control systems as well as technical barriers, the management could simply use computers or a single server, which have more computing power than the remaining participants of the network for the effortless override of internal controls. Moreover, if logged transaction can be altered or possible deleted ex post, transparency of financial information is useless. This leads to the aforementioned negative effects on the external control system, as retrospective fraud detection mechanisms that may be conducted through external auditors are getting ineffective and subsequently also large parts of investors or exchange market participants, which rely on disclosed and allegedly external audited financial information. Consequently, using PoW as DCM for blockchain-based accounting, MAF will not be impeded, neither from an organizational perspective by decentralization nor through immutability of data, i.e. the technical perspective. Notably, PoW would even ease the override of controls, as the management does not need to convince others to support and conceal the fraud such as in the case of Comroad. This especially holds in the absence of direct monetary incentives that encourage honest behavior such as in the case of Bitcoin.

4.2 Alternative Decentralized Consensus Mechanisms

Table 1 provides an overview over DCMs developed after the emergence of PoW. For the sake of completeness PoW-based consensus is also included. A differentiation of the DCMs is conducted according to their ability to allow for permission-less or permissioned access of nodes as well as whether the mechanisms facilitate decentralized validation or not. This differentiation is done in accordance to the second layer of the model presented in Figure 3.

Table 1. Overview of decentralized consensus mechanisms after the emergence of PoW

	Permission-less Access	Permissioned Access
Decentralized Validation	• Proof-of-Work • Proof-of-Stake • Proof-of-Work based derivatives • Federated Byzantines Agreement	• Proof-of-Work • Proof-of-Stake • Proof-of-Work based derivatives • Federated Byzantines Agreement
Centralized Validation	• Delegated Proof-of-Stake	• Redundant Byzantines Fault Tolerance • Ripple consensus • Bilateral node-to-node (N2N) • RAFT and derivatives • Delegated Proof-of-Stake

Certain DCMs enable both, permissioned as well as permission-less access (although they might designed to be used in a permission-less system at first). Contrarily, it is assumed that a mechanism, by intention, will not feature decentralized validation and centralized validation at the same time. However, it is acknowledged that in practice decentralized validation may exhibit centrality tendencies. In the following, DCMs will be analyzed, if they feature permissioned access as well as decentralized validation and are, therefore, suited for an application according to the preciously presented scenario of blockchain-based accounting.

Proof-of-Stake

Proof-of-Stake (PoS) is based on the assumption that PoW's dependence on energy consumption creates unnecessary cost overhead in networks. PoS is a form of proof of ownership of the currency in the network [19]. Instead of using relative hash rates of miners for network stability, the protocol splits blocks and the according transactions proportionally to the current wealth of miners [26]. In a blockchain-based accounting system stakes will most likely be stocks. As the management will probably have the majority of stocks (this was also observed in the case of the Comroad scandal [15]), the management would be enabled to change financial transaction registers at their will, without having to respect any control system. Thus, despite the Proof-of-Stake is initially designed to promote decentralized validation, in practice the validation of transaction using this mechanism will be centralized and most likely led by the management. Moreover the protocol exhibits other general security issued, such as the so-called "nothing at stake" attack, where attackers can commit collateral as they can go back and rewrite history from a point where they still had stake [19].

Proof-of-Work Based Derivatives

Proof-of-Activity (PoA) is a combination of PoW and the PoS and described as one example of different PoW derivatives. Finding consensus by using PoA consists of the transformation of a pseudorandom value into a satoshi, which is the smallest unit of the cryptocurrency Bitcoin. According to [6], this is done by selecting a pseudorandom index between zero and the total number of satoshis in existence up to the last block, inspecting the block in which this satoshi was minted and following each transaction that subsequently transferred this satoshi to an address until reaching the address that currently controls this satoshi. Only active stakeholders, who maintain a node, get rewarded in exchange for the service they provide for the network. Despite PoA induces less overhead in terms of communication, it does not prevents "nothing-at-stake"-attacks [6] and therefore, does not guarantee for the fraud resistance of data on the blockchain. For the sake of completeness it should be mentioned that there exist further DCMs such as proof-of-capacity or proof-of-burn that are based on or are related to PoW. However, they are rather used for distributed payment systems and rarely discussed for other appliances in a scientific context.

Federated Byzantines Agreement

Federated Byzantines Agreement (FBA) allows each node to select a set of other trusted nodes, which induces so-called flexible trust, meaning that all users have the freedom to trust any combination of parties. Nodes may select those participants based on arbitrarily criteria such as repudiation. To find consensus, a node waits for the vast majority of trusted nodes (quorum slice set) to agree on a transaction before considering the transaction settled. In turn, those nodes do not agree to the transaction until the participants they consider as important agree to the transaction as well, and so on. The key distinction between the FBA and prior Byzantines Agreements is the individual and decentralized trust decisions. If enough network nodes accept a transaction, it becomes infeasible for an attacker to roll it back [23, 29]. Moreover, security rest on digital signatures and hash families whose parameters can realistically be tuned to protect against adversaries with unimaginably vase computing power [23].

Notably, the FBA is a majority voting system, related to the decisions of selected trusted nodes. As in every voting system and, especially, if voting nodes are known to each other in a closed system, strategic voting cannot be excluded. Thus, it may be easy for the management to couple votes of particular nodes and their influence on other nodes to career perspectives and/or monetary or non-monetary incentives, leading to a strong influence of the management on the voting outcome and data that will be incorporated on the blockchain. Moreover, the management may also be able to influence the majority of nodes to subsequently alter data on the blockchain to cover up traces. Consequently, FBA is not able to prevent the occurrence of MAF.

5 Conclusion and Outlook

Academic and industrial work proposing the application of blockchain technology for accounting emphasize the immutability of financial recording based on a proof-of-work based decentralized consensus, probably leading to fraud-resistance. After identifying one of today's core problems of proper accounting, i.e. MAF, and proposing a scenario for blockchain-based accounting within a public company, this paper asserts that PoW is not effective in terms of preventing MAF. This conclusion is based on the assumption that there exist no incentive that prevents the management from committing fraud in accordance to [9, 16, 32], rather it is the mechanism, here distributed consensus that probably could prevent the commitment of MAF. Moreover, proof-of-work is even expected to ease the conduction and concealment of MAF, owing to the prevailing centrality tendencies within the system. Alternative decentralized consensus protocols were examined in accordance to the presented scenario of blockchain-based accounting. This paper concludes that currently, there exist no DCMs that promote permissioned systems, featuring decentralized validation and simultaneously preventing MAF. Overall, the ability of the blockchain and DCMs in the proposed scenario might be overestimated or even overhyped, even if a certain general potential of the technology in accounting could be attested owing to its decentralized and transparent nature. However, proposals for concrete applications must be strongly oriented on the de facto problems such as in the case of accounting and MAF. Accordingly, further research should focus on the development of advanced consensus mechanisms that take into account the above-discussed issues, and especially, the ability of management override of controls. Variations in the proposed scenario are also conceivable. Overall, a special emphasis should lie on the cost-efficiency of such systems as well as security as basic requirements. Without these two prerequisites it is hard to imagine that any such system will be implemented in the future. Moreover, a comparison of intra- and inter-corporate solution as well as other possible scenarios should be pursued, in order to receive a more compressive evaluation of the potential of blockchain-based accounting.

References

1. Adams, R.B., Ferreira, D.: A theory of friendly boards. J. Finance. 62, 1, 217–250 (2007).
2. Andersen, N.: Blockchain Technology A game-changer in accounting? (2016).
3. Beasley, M.S.: An empirical analysis of the relation between the board of director composition and financial statement fraud. Account. Rev. 71, 4, 443–465 (1996).
4. Beasley, M.S. et al.: Fraudulent Financial Reporting. Committee. 12, 60 (2010).
5. Becker, J. et al.: Can we afford integrity by proof-of-work? scenarios inspired by the bitcoin currency. In: The Economics of Information Security and Privacy. pp. 135–156 (2013).
6. Bentov, I. et al.: Proof of Activity: Extending Bitcoin’s Proof of Work via Proof of Stake. (2014).
7. Bungartz, O.: Handbuch Interne Kontrollsysteme (IKS): Überwachung und Steuerung von Unternehmen. Erich Schmidt Verlag, Berlin (2012).
8. Byström, H.: Blockchains, Real-Time Accounting and the Future of Credit Risk Modeling. (2016).
9. Caplan, D.: Internal Controls and the Detection of Management Fraud. J. Account. Res. 37, 1, 101–117 (1999).
10. Daum, R.: Phantompartner und Phantasieumsätze in Asien - Enttarnung eines Börsenstars. In: Leif, T. (ed.) Mehr Leidenschaft Recherche: Skandal-Geschichten und Enthüllungsberichte - Ein Handbuch zur Recherche und Informationsbeschaffung. pp. 134–143 Westdeutscher Verlag, Wiesbaden (2003).
11. Daum, R.: Phantompartnern in Asien auf der Spur. In: Schröder, C. and Sethe, R. (eds.) Kapitalmarktrecht und Pressefreiheit. pp. 9–30 Nomos (2011).
12. Decker, C., Wattenhofer, R.: Information propagation in the Bitcoin network. In: 13th IEEE International Conference on Peer-to-Peer Computing, IEEE P2P 2013 - Proceedings. (2013).
13. Deutschland, B.: Aktiengesetz (AktG), (2015).
14. Dinger, J., Hartenstein, H.: Die vermeintliche Robustheit von Peer-to-Peer-Netzen. (2006).
15. Dorin, M.: Institutionelle Maßnahmen zur Verbesserung der Qualität von Abschlußprüfung. Bielefeld University (2006).
16. Feng, M. et al.: Why do CFOs become involved in material accounting manipulations?

J. Account. Econ. 51, 1–2, 21–36 (2011).
17. Harrast, S.A., Mason-Olsen, L.: Can Audit Committees Prevent Management Fraud? CPA. 77, 1, 24–27 (2007).
18. IDW: IDW PS 261: Feststellung und Beurteilung von Fehlerrisiken und Reaktionen des Abschlussprüfers auf die beurteilten Fehlerrisiken. WPg Suppl. 2, (2012).
19. King, S., Nadal, S.: PPCoin: Peer-to-Peer Crypto-Currency with Proof-of-Stake. Ppcoin.Org. (2012).
20. Kuizick, R.S.: Sarbanes-Oxley: Effects on Financial Transparency. SAM Adv. Manag. J. 69, 1, 43–49 (2004).
21. Lamport, L. et al.: The Byzantine Generals Problem. ACM Trans. Program. Lang. Syst. 4, 3, 382–401 (1982).
22. Lennox, C., Pittman, J.A.: Big five audits and accounting fraud. Contemp. Account. Res. 27, 1, 209–247 (2010).
23. Mazières, D.: The Stellar Consensus Protocol: A Federated Model for Internet-Level Consensus. (2016).
24. Nakamoto, S.: Bitcoin: A Peer-to-Peer Electronic Cash System. Consulted. 1–9 (2008).
25. Peters, G.W., Panayi, E.: Understanding Modern Banking Ledgers through Blockchain Technologies: Future of Transaction Processing and Smart Contracts on the Internet of Money. arXiv Prepr. arXiv1511.05740. 1–33 (2015).
26. Pilkington, M.: Blockchain Technology: Principles and Applications. In: Olleros, F.X. and Zhegu, M. (eds.) Research Handbook on Digital Transformations. Edward Elgar, Cheltenham, UK (2016).
27. Probst, L. et al.: Blockchain Applications & Services. (2016).
28. Reason, J.: The contribution of latent human failures to the breakdown of complex systems. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 327, 1241, 475–484 (1990).
29. Rubin, J.: Federated Systems. , Massachusetts (2015).
30. Sawyer, L.B. et al.: Sawyer’s Internal Auditing: The Practice of Modern Internal Auditing. Institute of Internal Auditors (2012).
31. Swanson, T.: Consensus-as-a-Service: a brief report on the emergence of permissioned, distributed ledger systems. (2015).
32. Topgos, M.A.: Why Management Fraud Is Unstoppable: Certified Public Accountant. CPA J. 12, 34, 34–41 (2002).
33. Tsai, W.-T. et al.: A System View of Financial Blockchains. In: IEEE Symposium on Service-Oriented System Engineering (SOSE). pp. 450–457 , Oxford.
34. Van de Velde, J. et al.: Blockchain in Capital Markets: The Prize and the Journey. (2016).
35. Vukolic, M.: The Quest for Scalable Blockchain Fabric: Proof-of-Work vs. BFT Replication. In: International Workshop on Open Problems in Network Security. pp. 112–125 Springer International Publishing, Zürich, Switzerland (2016).
36. Yermack, D.: Corporate Governance and Blockchains, http://www.nber.org/papers/w21802, (2015).
37. Factom - Apollo, https://www.factom.com.
38. https://www.scorechain.com.

Towards a Smart Services Enabling Information Architecture for Installed Base Management in Manufacturing

Sonja Dreyer[1], Daniel Olivotti[1], Benedikt Lebek[2], and Michael H. Breitner[1]

[1] Leibniz Universität Hannover, Information Systems Institute, Hannover, Germany
{dreyer,olivotti,breitner}@iwi.uni-hannover.de
[2] bhn Dienstleistungs GmbH & Co. KG, Aerzen, Germany
lebek.benedikt@bhn-services.com

Abstract. In the manufacturing industry the provision of smart services is an opportunity to gain a competitive advantage. As there are high demands on machine availability, smart services in the field of installed base management are important. Through integrating condition monitoring data with installed base data a digital twin of the installed base can be created. This enables comprehensive analyses and the provision of individualized smart services. But this requires to structure and standardize the data. Following the action design research (ADR) approach, in this article design principles of an information architecture are developed. The architecture is evaluated and improved in the context of an international engineering and manufacturing company. A test run with real machine data shows the applicability in practice.

Keywords: Digital twin, information architecture, installed base management, smart services, product-service system

1 Introduction

In manufacturing industry, machinery and equipment are subject to high demands on availability and productivity [1], [2]. To meet these requirements, original equipment manufacturers (OEMs) as well as machine component suppliers shift in emphasis from selling only technical products to offering additional individualized smart services [3], comprising e. g. maintenance, repair, spare parts delivery, process consulting. Services are an economic guiding force for OEMs and machine component suppliers [4] because the supply of services in addition to products leads to new sales opportunities and to greater customer loyalty [5-7].

When considering individualized smart services, a major challenge for OEMs and machine component suppliers is that they often lack knowledge about the state of the machinery and equipment during the use phase [1]. Therefore, the collection and processing of condition monitoring data (field data) has been discussed within academic literature as inevitable for offering guaranteed machinery and equipment availability and productivity [8], [9]. However, the integration of condition monitoring

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Dreyer, S.; Olivotti, D.; Lebek, B.; Breitner, M.H. (2017): Towards a Smart Services Enabling Information Architecture for Installed Base Management in Manufacturing, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 31-45

data with installed base data (i.e. product master data, service data and contractual data [10]) is a necessary precondition for the provision of individualized smart services that has been rarely recognized by researchers. Organizations in value networks (i.e. OEMs, machine component suppliers and machine owners/operators) face the problem of how to use existing data from IT systems (e. g. ERP, MES, CRM) to depict machines with the respective installed components, locations, maintenance protocols etc. and how to enrich them with real-time data. Creating a digital twin based on these data provides the basis for rendering individualized smart services in value networks. Digital twins are realistic models representing machines with all their components, their current state as well as their interaction with the environment [11], [12]. Therefore, a digital twin is not static but dynamic which is why a standardized and consistent data and information management is necessary [13]. Although researchers began to work out information sources for installed base characterization [14], a generalized information architecture for installed base management that aims to enable smart services is still missing. An information architecture enables to store and process data, to analyze and evaluate them and to use them to offer services. Therefore we proposed the following research question:

> *RQ: What are general design principles of an information architecture for installed base management that enables smart services?*

To answer this research question and to build a bridge between academic rigor and practical relevance, we adapted the action design research (ADR) approach [15] in the context of an international engineering and manufacturing company. With the use of different cycles, ADR allows continuous interaction between researchers and practitioners in early stages.

The remainder of this article is structured as follows: Section 2 introduces the action design research approach and its application. The information architecture is designed in section 3, including a literature review in the field of installed base management. General design principles are carried out and results are discussed in chapter 4. The article ends with limitations and conclusions in sections 5 and 6.

2 Research Design

Starting from existing work in the research field of installed base management, the aim of this study was to develop an information architecture for practitioners to enable individualized smart services. Additionally, design principles of an information architecture for installed base management that can be applied within multiple companies were developed as theoretical contribution. For this purpose, the action design research (ADR) approach by Sein et al. [15] was selected as the underlying research methodology. Motivated by an increasing debate about the gap between organizational relevance and methodological rigor [16], [17] in IS research, ADR was developed in order to close this gap by presenting an integrative research approach of action research and design research [15]. ADR is well suited for deriving generally applicable knowledge while incorporating two main challenges. First, in order to generate practically relevant outcomes, ADR aims to solve a problem in an organizational setting by facilitating ongoing interaction of researchers and practitioners. Second, in order to account for academic rigor, ADR develops generalized design principles that address a class of problems through formalized learning from organizational intervention (Figure 1).

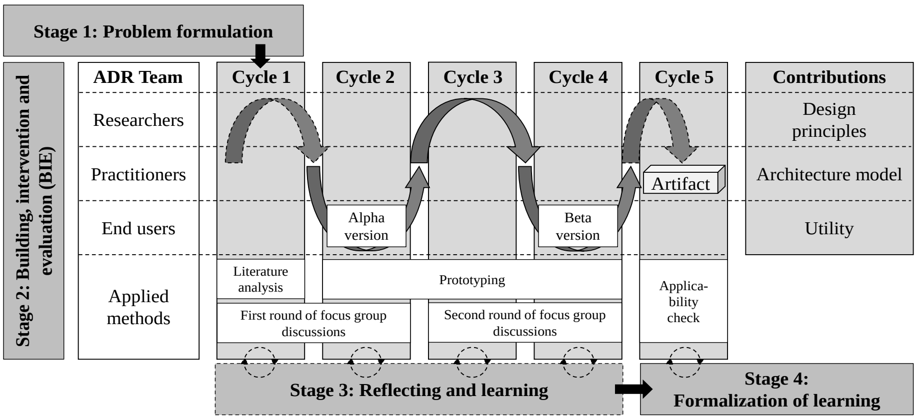


Figure 1. Research design based on the ADR approach from Sein et al. [15]

In the context of this study, the first stage was driven by a problem perceived at an international engineering and manufacturing company. In order to offer individual smart services in the area of installed base management, the target company required an information architecture to be able to organize and analyze machine owner's installed base data. In stage one the specific practical problem was formulated as an instance of a broader class of problems. An ADR team was formed, made up of researchers from a German university, practitioners from the target company's IT department and product service account management as well as end users from customer service department. The shared competencies facilitated the problem definition and formulation. In the building, intervention and evaluation (BIE) stage, five iterative cycles were carried out in order to build and continuously evaluate a prototype for an information architecture for installed base management. In the first cycle of the BIE stage, an initial prototype ('Alpha version') was derived from requirements analysis that was conducted during the problem formulation stage, as well as analysis of existing literature and a series of focus group discussions involving practitioners and end users. Subsequently, the alpha version was presented to practitioners and end users for evaluation by conducting a second series of focus group discussions in cycle two. Each focus group discussion was documented and qualitatively evaluated afterwards [18]. Based on the feedback gathered in the previous cycle, the prototype was further improved and specified ('Beta version') in cycle three. The applicability of the proposed information architecture was assessed during a test run with real data from a customer organization in cycle four. For the purpose of evaluation, next to end users, members of the target company's service product management, service account management, IT application development and IT strategy departments were involved. Feedback was used for incremental reshaping of the prototype in cycle five until the final version was reached. For continuous evaluation of each development step of the proposed information architecture, stage three (reflection and learning) was carried out simultaneously to the BIE stage. This procedure allowed a better understanding of the problem from different points of view. In stage four, learning from the specific solution was generalized in order to address a broader class of problems. Hence, the developed information architecture aims for general applicability for various companies and is not limited to the use case of the target organization. The development of design principles contributes to this.

3 Information Architecture Development

3.1 Problem Formulation

The presented information architecture emerged from a project of an international engineering and manufacturing company for storing and structuring machine owner's installed base data. Agile project management was used to realize the project what is appropriate to the ADR approach. The executing company operates in 60 countries and has the headquarters in Germany. The project concerned a service where machine and component data of a customer's machine are collected and structured subsequently. A detailed overview of the built-in components and its condition is given through creating a digital twin. This provides the basis to offer individualized smart services.

A use case was defined as an example for smart services in the manufacturing industry which are basing on installed base management. The examples arose from focus group discussions with employees of a large German automotive manufacturer, a machine owner. The participants of the focus group discussion came from the departments of maintenance processes, systems technology and automation technology. Additionally, practitioners from the engineering and manufacturing company, a component supplier, formed part of the group. They mainly came from the IT and the services department. The discussion dealt with services that would help to simplify tasks related to asset management. Mainly, there were six key questions the participants of the focus group discussion have which are not answered until now.

Table 1. Key questions of a machine owner regarding installed base management

Key question
How can installation errors be avoided?
How can error causes be identified immediately and reliably?
How can maintenance efforts be minimized?
How can maintenance schedules be planned optimally?
How can we learn from experiences?
How can knowledge be provided at the right time and at the right place?

All the questions, shown in Table 1, were answered within a use case for smart services. The use case as an example for smart services contributes to specify the information architecture for installed base management. The idea of the use case was to use sensor data and to connect them with further installed base data to identify the current state of the machine as well as of the integrated components. When knowing exactly for what kind of task the components are used analyses enable to identify error causes immediately and reliably. During the production, it does not only have to be looked at a single machine but the context of all machines in a production line is important. For example, on the basis of digital twins maintenance schedules can be optimized. This includes systematic maintenance as well as predictive maintenance depending on the analyses results of the data. Such a smart service would help the maintenance department to ensure a high availability of the machines. Additionally, an integrated visualization cockpit enables to support the machine operator. In case of an error the machine operator receives concrete instructions what has to be done to solve the problem. This requires that a history of former errors, the respective error code and solutions exist. In a first step, an information architecture was necessary to be able to structure, standardize and analyze installed base data. Requirements for the architecture were determined in several focus group discussions as well as from existing literature.

3.2 Related Literature

According to the first cycle of ADR, a comprehensive literature review was conducted [19] to identify and analyze the current state of research in the area of installed base management and digital twins. A list of search terms was predefined and it was mainly searched through seven databases: AISeL, IEEExplore, JSTOR, ResearchGate, Science Direct, SpringerLink, Taylor & Francis.

Table 2. Number of hits for different search terms

Search term	*Hits*	*Temporal and thematic restriction*	*Relevant title and abstract*	*Strong relation to research topic*
“installed base management“	84	22	92	26
“asset management”	20389	1042		
“inventory management”	4090	706		
“inventory data”	14957	382		
“asset data”	1715	585		
“machine data”	3645	711		
“installed base data”	25	24		
“digital twin”	53	53		
“digital shadow”	71	71		

As there were many hits due to the search terms (Table 2), the hits were temporally restricted. Publications older than 1995 were excluded. Furthermore thematically restrictions were made and only publications in journals and proceedings in fields related to information systems were included. The identified articles were checked for general relevance for the research topic by title and abstract. Afterwards the remaining publications were examined regarding their contribution to at least one of six categories developed (Table 3) by considering the whole article. The term installed base management is rarely discussed in academic literature. Asset management has a close relation to installed base management which is why this is also part of the literature review. According to Lin et al. [8] the objective of asset management is to support and optimize the lifecycle of physical assets. As an extension to asset management, in installed base management the individual components of the machines and their interplay within the machine and across them are also in focus. Installed base management goes deeper into the machines to ensure a high machine availability. But in contrast to installed base management asset management has found plenty of recognition by researchers.

Table 3. Literature categorized by focus

Author(s)	*Data sources*	*Data quality*	*Using large datasets*	*Digital twin*	*Information architecture*	*Services*
Abramovici et al. 2016 [27]	●	●		●		○
Bahga, Madisetti 2012 [25]			●		○	
Borchers, Karandikar 2006 [14]	●	●				○
Cai, Ziad 2003 [23]	○	●	○			
Felden, Buder 2011 [35]	●	○			●	
Furtak et al. 2015 [26]		●	●			●
Jun, Fang 2013 [36]		○		●		
Gabor et al. 2016 [12]	○			●	●	●
Lee et al. 2013 [13]			●	●	●	○
Lim et al. 2015 [37]						●
Luoto 2013 [21]	○	●				
Mert et al. 2016 [1]						●
Mohseni 2003 [30]	○				○	●
Mueller et al. 2003 [32]					●	
Narayanamurthy, Arora 2008 [38]					●	○
Neely 2008 [29]						●
Neff et al. 2012 [31]					○	●
Power 2014 [24]	●		●			
Rosen et al. 2015 [11]	●			●		○
Sarfi et al. 2012 [22]	○	●	○			●
Thies, Stanoevska-Slabeva 2011 [39]	○	●			●	●
Turunen, Toivonen 2011 [20]	●					●
Wollschlaeger et al. 2015 [33]	●				●	
Zampou et al. 2015 [34]	●	●			●	
Ziekow et al. 2010 [40]	○				●	
Zolnowski et al. 2011 [28]						●

● Considered topic ○ Partially considered topic

Borchers et al. [14] outlined that the sources of installed base data are manifold. Sales and enterprise resource planning (ERP) systems, production databases and service information systems are the sources mentioned by the authors. Turunen et al. [20] supplemented this list and indicated that field technicians and other operational staff are an important channel for additional information. Louto [21] agreed and noted that manually collected data plays a decisive role when executing services. Sarfi et al. [22] consider data quality as an important aspect in the field of asset management. Cai et al. [23] looked more closely at the completeness of the data as part of data quality and proposed a method for evaluating it. Because a huge amount of data is required to maintain a good record, Power [24] described how to use semi-structured and unstructured large datasets for analytics. Although the author emphasized machine data as a main contributor to adding value, the article gives a more general overview of the topic of large datasets. Bahga et al. [25] looked especially at machine data and how to analyze it. But in their article, machine data is put on a level with sensor data. The concentration on sensor technology is pursued by Furtak et al. [26]. Creating a digital twin, also named as digital shadow, is rarely considered in academic literature. Abramovici et al. 2016 [27] define digital twins as virtual smart products. Gabor et al. [12] see digital twins as the basis for further applications. Rosen et al. 2015 [11] agree to that and discuss the importance of digital twins for autonomous systems in manufacturing. They summarized that digital twins are important in the whole lifecycle of a product [11]. As there is no additional value unless the data is used to generate new services, the service aspect is subject of discussion in several publications. Services based on remote technology is covered by Zolnowski et al. [28]. Neely [29] agreed that collecting data and analyzing it forms the basis for implementing new services in the area of asset management. Not only do business models have to be created, the implementation of an IT solution becomes necessary for using a data management platform for services. Mohseni et al. [30] stated that a management tool that integrates different aspects is necessary to manage an asset successfully. Müller et al. [32] explained how to realize access to field device information by means of an information server. A reference architecture for condition monitoring was presented by Wollschlaeger et al. [33]. Zampou et al. [34] deduced an architectural framework for environmental performance monitoring of machines.

Results indicated that many publications emphasize the importance of generating services out of installed base management (see Table 3). The existence of different data sources are also part of many publications, in contrast to the topic of creating a digital twin. Information architectures for asset management were also a research topic. But researchers are mainly focusing on condition monitoring data to offer services. Using existing data to describe components, machines and whole plants and enriching them with real-time data is not the focus of research. Therefore, to the best of our knowledge, there is no information architecture existing which is meeting the requirements resulting from the project goal. An information architecture for installed base management to enable smart services is still missing.

3.3 Requirements on an Information Architecture

A series of focus group discussions with different employees from an international engineering and manufacturing company was conducted to ensure creating an information architecture for installed base management that is relevant in practice. The participants came from the fields of business process consulting, application development, product management, service account management, and technical customer service. For a better structure, the discussion was divided into three parts: objective of the information architecture, general structure, and information architecture design.

The focus group stated that the objective is to be able to structure and standardize installed base data. Organizing the data into structures is important to simplify its further use [33]. Aggregating data from different sources should be enabled to create a basis for installed base management because using existing data prevents duplication of data. This means that existing data should be used and imported as required. Existing data, manually captured data and sensor data should be brought together in order to create a digital twin. To simplify subsequent analyses a uniform structure of components and machines should be ensured. The focus group recommended referring to the superior object of a component by creating a hierarchical structure. A structure from general to specific is named as suitable to represent a machine and its components.

Data quality is an important factor in installed base management [22]. The quality of data has to be ensured for successful asset management and therefore for installed base management [8]. It is mentioned by the focus group that the applicability of the information architecture should not be limited to the German-speaking area. General formats should be found because engineering and manufacturing companies often operate worldwide. The possibility of using a single information architecture for installed base management to enable smart services contributes to avoid isolated solutions. Regarding analyses of the data participants of the focus group mentioned that it should be possible to analyze structured as well as unstructured data like free texts. It is stated that sometimes it is not possible to find a uniform format, e. g. to describe the current condition of a machine. Analyses across machines and plants should also be possible. Diverse access rights for different persons would contribute to transparency.

3.4 Information Architecture Design

With focus on the requirements determined a conceptual model of the information architecture was designed (Figure 2). The first layer describes the data storage and processing. Many different types of data have to be considered to be able to reach the objective of a digital twin. For example, data generated during the production phase of a component is important. Another data pool that was discussed by the focus group was the topic of successor data. The requirement is guaranteeing that only current products can be sold. The orientation on the product lifecycle helps to meet this requirement. Another source of data considered is the installed base itself. This data is mainly collected on-site by service technicians and includes the installed base structure, a condition description of the whole machine and a description of the individual components. As required by the focus group condition monitoring data such as sensor data were also considered. The data is stored in different systems, depicted due to a systems sub-layer. All data saved in different systems is transferred to a database management system to relate the data for a component or machine. The second layer concerns services. They are necessary to generate value based on the digital twins created. On the one hand the data of components, machines and plants have to be analyzed. It is also possible to compare the data across components, machines, or even plants. On the other hand a knowledge management is necessary to turn data into information. E. g. when sensor data indicated an anomaly, knowledge is necessary to interpret the data and to give advices to the machine operator. Tools such as mathematical models may also help evaluate the analysis results. The third layer concerns the presentation and is the interface for the participants of the value network. Role-based user interfaces enable to individually define read and write permissions.

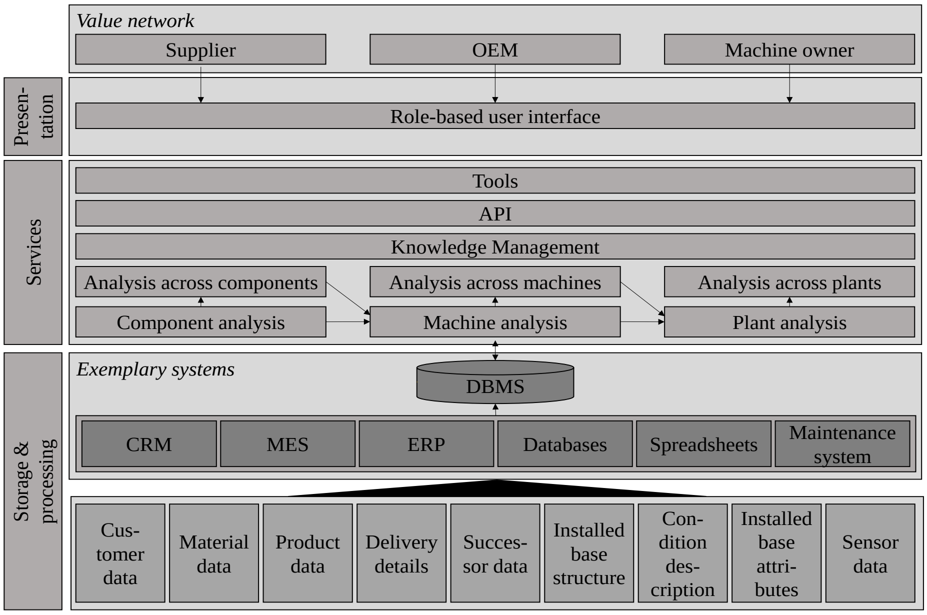


Figure 2. Conceptual model of the information architecture

3.5 Evaluation and Optimization

The implemented information architecture was evaluated in a test run. Real installed base data collected in a production site was organized through the architecture to identify optimization possibilities. Proving a conceptual model in practice leads to great improvements [36]. With the test run, it was ensured that the platform is applicable in practice as an applicability check is an essential part of a research process [41]. Additionally, focus group sessions as part of the evaluation helped to further improve the information architecture [16]. Therefore, a series of focus group discussions with participants from the application development, product management, service account management, and technical customer service of an engineering and manufacturing company were conducted to find further points for improvement.

With basis on the information architecture it is possible to create a digital twin of the machine. The focus group appreciated the fact that existing data from other databases can be added to the respective component. For example, information on production already stored in a system can be assigned to the respective component. But it has to be ensured that the information imported is also understandable for external groups. The use case that was worked out as an example for smart services included the visualization of analyses results and further information about the machine, its current state and its environment. It is also possible to partially import data and to complement it with manually entered data. All predefined installed base data could be implemented and standardized. Documents and photos could also be attached to components. It is mentioned that it should be easy to integrate new databases or spreadsheets. When using the serial number to insert existing data, it should be ensured that this data is reliable and kept up to date. Furthermore, it should be ensured that the data storage and structure is consistent for the different components and machines. This is necessary to automatically analyze the data for offering smart services subsequently.

In accordance with the fifth cycle of ADR, the points of criticism were taken up to optimize the information architecture and to create the final artifact. The valuable recommendations led to improvements of the information architecture. Thus, the evaluation helped to further simplify the creation of a digital twin of machines and whole plants with the objective to offer smart services for machine owners.

4 Formalization of Learning and Discussion of Results

Following the ADR approach, an information architecture to manage installed base data was developed. The five cycles presented in this research method were used to create an artifact that is relevant in practice. Several focus group discussions and a test run were conducted to ensure practical relevance. Through formalization, the learning was converted into general design principles to contribute academic knowledge to the research field (Table 4). At the beginning of the project, transparency was named as a challenge to create a digital twin of a machine. This means that it has to be clear how the structure of the digital twin is designed. It ensures that the structure is understood by different persons and also facilitates analyses. As it should be possible to identify the components, they have to be named uniformly. Nevertheless, an unequivocal number for each component should exist to enable locating a specific component. In the presented project, each component had a serial number for clear identification. For example, this is important when a component has to be replaced because it is affected by a production error. Standardized data is a precondition for further analyses and evaluation. For the same information it should always be used the same data format. The structure as well as information does not only have to be understood by employees of only one apartment or a single plant but in the whole company what requires international valid regulations to avoid misunderstanding. The participants of the focus group discussions that were asked regarding their requirements towards an information architecture worked at a component supplier. But as there are several participants in the value network apart from component suppliers, different perspectives on the digital twin are necessary. For example, machine owners are primarily interested in analyses concerning their own machines and plants. A role-based authentication is the tool to realize this. All named design principles contribute to an easy analysis and evaluation of the collected data across components, machines and plants. It is depicted in the services layer in the developed information architecture. As it is not possible to structure all data, for example because of free texts from employees in the field, semantic analyses should also be enabled. A service orientation of the information architecture ensures that tools such as mathematical models can be integrated. Real-time data handling is inevitable to be able to store and process sensor data.

Table 4. Set of design principles

Design principle	*Description*	*Examples*	*Standards/ best practices*
Transparency - Consistent vocabulary - Clear allocation of components - Clear identification of products	It is necessary to have a clear hierarchical structure of the data. The naming should be consistent and generally comprehensible. An unequivocal identification of the products contributes to the clear structure and the creation of a digital twin.	Uniformly named components, serial numbers for unequivocal identification of components	Extensible Markup Language (XML), International Standard Serial Number (ISSN)
Standardization - Uniform data format - Machine readability of the data	A uniform format of the data is necessary for further analyzing. It ensures both that the data is understandable for different target groups and readable by machines.	Uniform sensor data format, enabling exchangeability between companies	eCl@ss
Internationality - International data format - Transferability to other languages	As organizations in the manufacturing industry often operate worldwide, it is of importance that the data has an internationally understandable format. The transferability to other languages should be given.	Uniform date format, multiple language data maintenance	ISO (e. g. date/time ISO 8601)
Perspectives - Adaptable structure depth - Adaptable access rights	The data is used by different participants and by users of smart services. Therefore, a role-based authentication with different read and write permissions is required.	User-dependent view, selective transaction authorization	One-time passwords (OTP), Certified-based Authentication (CBA)
Analysis - Across components, machines, plants - Unstructured data	It should be possible to analyze the data, independent on whether they are structured or not.	Comparing the state of different machines, analyzing unstructured comments in text boxes	Apache Hadoop
Service orientation - Tool integration - (near) Real-time data handling	Tools enable individualized smart services. The use of condition monitoring data for smart services requires real-time data handling.	Visualization tools, automated reports, real-time sensor data processing	

The information architecture enables to offer smart services through analyzing and evaluating installed base data. From a supplier's or an OEM's perspective it becomes possible to see the machine owner's needs in the short and medium term. Building on this, further services can be offered. The chance of participating in future projects is higher because of the direct contact with the machine owner. As a consequence, sales and customer satisfaction as well as loyalty are increasing. The consequence is a necessary service orientation because it is only possible to generate value out of the analysis results when offering services. For example, in the described use case visualizing data and information was important. In the information architecture this is represented due to the presentation layer. Tools may be necessary to realize individualized smart services that are carried out automatically. As results from condition monitoring data such as sensor data are an important part of the digital twin, real-time data handling is necessary. In the information architecture the data storage and processing layer represents this.

5 Limitations and Further Research

The created information architecture provides the opportunity to manage installed base data efficiently. However, certain limitations have to be considered. The information architecture was designed in a project of an engineering and manufacturing company. Installed base data collected by service technicians and existing data of the company were used as basis for defining the necessary data types. This was supplemented by requirements stated during a series of focus group discussions in the same company. When talking about an installed base management different participants of the value network have to be taken into account. In the presented project it is focused on a component supplier and its requirements. Therefore, it cannot be excluded that the architecture has to be adapted for other participants. In future research the requirements of other participants of the value network will be focused. Real-time data was no part of the test run. The company did not provide real-time condition monitoring data, thus it was not possible to perform a test run for this kind of data. Therefore, it cannot be said whether adjustments are necessary to be able to handle this high amount of data.

As the architecture was developed and tested in the headquarters in Germany, it is not sure whether the transfer to other languages is possible without much effort. The data were mainly stored and structured in German and a transfer to other languages was not necessary in the test run. The information architecture provides the opportunity to unite different data types and to standardize them. Smart services based on installed base management can be provided. Depending on the offered service, it could be necessary to expand the architecture. Using the information architecture for other smart services apart from the proposed use case will be tested in the future. Evaluating the information architecture flexibility and extensibility is another interesting approach for future research. Similar restrictions regarding the adaptability apply to the installed base data. It is not certain whether it is possible to standardize all data types in the future because future requirements are not yet known. Restrictions also apply to the evaluation of the information architecture. The information architecture was evaluated in a test run and not over a longer period in practice. Therefore, the information architecture was not checked under real conditions. The evaluation with other data may also lead to further enhancements and optimizations. Furthermore, it would be interesting to evaluate the information architecture and the developed design principles in focus group discussions with employees from different participants of the value network.

6 Conclusions

In this article, an information architecture to offer smart services based on installed base management in the engineering and manufacturing industry was worked out. Following the ADR approach, requirements on an information architecture for an engineering and manufacturing company were developed in a series of focus group discussions and due to a literature review. Structuring and standardizing installed base data to be able to analyze and evaluate them afterwards is inevitable to provide smart services. As an example a use case for a smart service was worked out. Based on this, the information architecture was created and refined. This was also done through several focus group discussions. Considering the discussion and limitations, the information architecture contributes to theoretical and practical knowledge. From an organization’s perspective, the information architecture enables to create a digital twin to offer individualized smart services. Answering the research question, formalization of learning led to design principles which contribute to academic knowledge in this area. Both the developed information architecture and design principles for creating an information architecture provide the basis for future research and expanding and optimizing the digital twin of machine owner’s installed base.

References

1. Mert, G., Herder, C.F., Menck, N., Aurich, J.C.: Innovative Services for Customized, Availability-oriented Business Models for the Capital Goods Industry. Procedia CIRP 47, pp. 501-506 (2016)
2. Haider, A.: Information Technologies for Engineering Asset Management - Cultural and Technical Barriers. In: Proceedings of the International Conference on Information Resources Management, Paper 4, Seoul (2011)
3. Schrödl, H., Bensch, S.: E-Procurement of Cloud-based Information Systems - a Product-Service System Approach. In: Proceedings of the 34th International Conference on Information Systems, Milan (2013)
4. Rai, A., Sambamurthy, V.: Editorial Notes - The Growth of Interest in Services Management: Opportunities for Information Systems Scholars. Information Systems Research 17 (4), 327-331 (2006)
5. Oliva, R., Kallenberg, R.: Managing the Transition from Products to Services. International Journal of Service Industry Management 14 (2), 160-172 (2003)
6. Cohen, M.A.: Product Performance Based Business Models: A Service Based Perspective. In: Proceedings of the 45th Hawaii International Conference on System Sciences, pp. 4814-4819, Hawaii (2012)

7. Barrett, M., Davidson, E., Prabhu, J., Vargo, S.L.: Service Innovation in the Digital Age: Key Contributions and Future Directions. MIS Quarterly 39 (1), 135-154 (2015)
8. Lin, S., Gao, J., Koronios, A.: Validating a Data Quality Framework in Engineering Asset Management. In: Proceedings of the 17th Australasian Conference on Information Systems, Paper 75, Adelaide (2006)
9. Fellmann, M., Hucke, S., Breitschwerdt, R., Thomas, O., Blinn, N., Schlicker, M.: Supporting Technical Customer Services with Mobile Devices: Towards an Integrated Information System Architecture. In: Proceedings of the 17th Americas Conference on Information Systems, Paper 250, Detroit (2011)
10. Jalil, M.N., Zuidwijk, R.A., Fleischmann, M., van Nunen, J.A.E.E.: Spare Parts Logistics and Installed Base Information. Journal of the Operational Research Society 62 (3), 442-457 (2011)
11. Rosen, R., von Wichert, G., Lo, G., Bettenhausen, K.D.: About the Importance of Autonomy and Digital Twins for the Future of Manufacturing. IFAC-PapersOnLine 48 (3), 567-572 (2015)
12. Gabor, T., Belzner, L., Kiermeier, M., Beck, M.T., Neitz, A.: A Simulation-Based Architecture for Smart Cyber-Physical Systems. In: Proceedings of the 2016 IEEE International Conference on Autonomic Computing, pp. 374-378, Würzburg (2016)
13. Lee, J., Kao, H.-A., Yang, S.: Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. In: Proceedings of the 6th CIRP Conference on Industrial Product-Service Systems, pp. 3-8, Windsor (2014)
14. Borchers, H.W., Karandikar, H.: A Data Warehouse Approach for Estimating and Characterizing the Installed Base of Industrial Products. In: Proceedings of the 2006 International Conference on Service Systems and Service Management, pp. 53-59, Troyes (2006)
15. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R.: Action Design Research. MIS Quarterly 35 (1), 37-56 (2011)
16. Lindgren, R., Henfridsson, O., Schultze, U.: Design Principles for Competence Management Systems: A Synthesis of an Action Research Study. MIS Quarterly 28 (3), 435-472 (2004)
17. Iivari, J.: A Paradigmatic Analysis of Information Systems As a Design Science. Scandinavian Journal of Information Systems 19 (2), 39-64 (2007)
18. Mayring, P.: Qualitative Inhaltsanalyse: Grundlagen und Techniken. Beltz, Weinheim and Basel (2015)
19. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26 (2), xiii-xxiii (2002)
20. Turunen, T.T., Toivonen, M.: Organizing Customer-oriented Service Business in Manufacturing. Operational Management Research 4 (4), 74-84 (2011)
21. Luoto, S.: Understanding the Aspects of Collecting Installed Base Information in Manufacturing Context - Towards a Future Research Agenda. In: Proceedings of the 2nd International Through-life Engineering Services Conference, pp. 416-419, Cranfield (2013)
22. Sarfi, R.J., Tao, M.K., Lyon, J.B., Simmins, J.J.: Data Quality as it Relates to Asset Management. In: Proceedings of the IEEE PES Transmission and Distribution Conference and Exposition, Orlando (2012)
23. Cai, Y., Ziad, M.: Evaluating Completeness of an Information Product. In: Proceedings of the 9th Americas Conference on Information Systems, pp. 2274-2281, Tampa (2003)
24. Power, D.J.: Using 'Big Data' for Analytics and Decision Support. Journal of Decision Systems 23 (2), 222-228 (2014)
25. Bahga, A., Madisetti, V.K.: Analyzing Massive Machine Maintenance Data in a Computing Cloud. IEEE Transactions on Parallel and Distributed Systems 23 (10), 1831-1843 (2012)

26. Furtak, S., Avital, M., Pedersen, R.U.: Sensing the Future: Designing Predictive Analytics with Sensor Technologies. In: Proceedings of the 23rd European Conference on Information Systems, Paper 54, Münster (2015)
27. Abramovici, M., Göbel, J.C., Dang H.B.: Semantic Data Management for the Development and Continuous Reconfiguration of Smart Products and Systems. CIRP Annals - Manufacturing Technology 65 (1), 185-188 (2016)
28. Zolnowski, A., Schmitt, A.K., Böhmann, T.: Understanding the Impact of Remote Service Technology on Service Business Models in Manufacturing: From Improving After-Sales Services to Building Service Ecosystems In: Proceedings of the 19th European Conference on Information Systems, Paper 109, Helsinki (2011)
29. Neely, A.: Exploring the Financial Consequences of the Servitization of Manufacturing. Operations Management Research 1 (2), 103-118 (2008)
30. Mohseni, M.: What Does Asset Management Mean to You?. In: Proceedings of the IEEE PES Transmission and Distribution Conference and Exposition, pp. 962-964, Dallas (2003)
31. Neff, A.A., Herz, T.P., Uebernickel, F., Brenner, W.: The Influence of Information Technology on Industrial Services in the Manufacturing Industry - A Literature Review and Future Research Directions. In: Proceedings of the Pacific Asia Conference on Information Systems 2012, Paper 57, Hochiminh City (2012)
32. Müller, J., Epple, U., Wollschlaeger, M., Diedrich, C.: An Efficient Information Server for Advanced Plant Asset Management. In: Proceedings of the IEEE International Conference on Emerging Technologies and Factory Automation, pp. 66-73, Lisbon (2003)
33. Wollschlaeger, M., Theurich, S., Winter, A., Lubnau, F., Paulitsch, C.: A Reference Architecture for Condition Monitoring In: Proceedings of the 2015 IEEE World Conference on Factory Communication Systems, pp. 1-8, Palma de Mallorca (2015)
34. Zampou, E., Pramatari, K., Mourtos, I.: Design of Environmental Performance Monitoring Systems in the Supply Chain: The Role of Interoperability. In: Proceedings of the 23rd European Conference on Information Systems, Paper 212, Münster (2015)
35. Felden, C., Buder, J.J.: Integrated Information Supply for Decision Support in Grid Companies. Business & Information Systems Engineering 4 (1), 15-29 (2012)
36. Jun, L., Fang, M.: The Design of Asset Management Service Platform in Universities Based on Cloud Computing Model. In: Proceedings of the International Conference on Mechatronic Sciences, Electric Engineering and Computer 2013, pp. 1649-1652, Shenyang (2013)
37. Lim, C.-H., Kim, M.-J., Heo, Kim, K.J.: A Conceptual Framework for Designing Informatics-based Services in Manufacturing Industries. In: Proceedings of the 7th Industrial Product-Service Systems Conference - PSS, industry transformation for sustainability and business, pp. 72-77 (2015)
38. Narayanamurthy, G., Arora, S.: An Integrated Maintenance and Asset Management System (IMAMS). In: Proceedings of the 2008 Integrated Communications, Navigation, and Surveillance Conference, Bethesda (2008)
39. Thies, H., Stanoevska-Slabeva, K.: Towards Inter-Organizational Environmental Information Systems for Sustainable Business Networks. In: Proceedings of the AMCIS 2011, Paper 325, Detroit (2011)
40. Ziekow, H., Ivantysynova, L., Guenther, O.: Architectural Patterns for RFID Applications in Manufacturing. In: Proceedings of the ECIS 2010, Paper 72, Pretoria (2010)
41. Rosemann, M., Vessey, I.: Toward Improving the Relevance of Information Systems Research to Practice: The Role of Applicability Checks. MIS Quarterly 32 (1), 1-22 (2008)

A Competency Model for "Industrie 4.0" Employees

Loina Prifti[1], Marlene Knigge[1], Harald Kienegger[1], and Helmut Krcmar[1]

[1] Technical University Munich, Chair of Information Systems, Munich, Germany
{prifti,marlene.knigge,harald.kienegger,krcmar}@in.tum.de

Abstract. This paper analyzes employee competencies for employees with higher education in Industry 4.0. An Industry 4.0 competency model based on a behavioral oriented approach concerning three variants, namely Information Systems, Information Technology and Engineering is developed by extending the SHL Universal Competency Framework through a structured literature review and focus groups with academic staff. The presented study contributes to research by providing a starting-point for further research regarding employee competencies for Industry 4.0. It contributes to practice as the provided competency model can be applied to Industry 4.0 job descriptions.

Keywords: Digital Transformation, Industry 4.0, Internet of Things, Competency, Competency Model.

1 Introduction

Recent technological developments, such as sensors, cyber-physical systems, the Internet of Things (IoT), or smart networks will influence each area of our life. This development is referred to as the fourth industrial revolution, also known as "Industrie 4.0" or "Industry 4.0" (I4.0). I4.0 is widely used in the international context, however we focus on the German concept of I4.0. In this context it approaches some of the challenges the world is facing today including the rise of resource and energy efficiency, production, demographic change etc. [1]. Further, it offers a huge potential especially for Germany as a global leader in the manufacturing industry [1]. Germany also possesses significant information technology (IT) know-how and competencies in automation, embedded systems or smart networks [1]. This offers the perfect prerequisites for Germany to become a leader in I4.0. "In essence, I4.0 will involve the technical integration of cyber-physical systems into manufacturing and logistics and the use of the Internet of Things and Services in industrial processes." [1]

I4.0 will influence our working environments significantly. E.g., it will change processes in purchase, production, manufacturing, sales or maintenance by including concepts as smart manufacturing, smart maintenance as well as a high degree of automation and integration in all enterprise processes [2]. It will have far-reaching implications on business value creation, business models, downstream services, and work organization [1]. As a consequence, employees will be confronted with transformed work processes and business models as well as with new technologies [2]. The model of work organization will transform due to the disruptive nature of emerging

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Prifti, L.; Knigge, M.; Kienegger, H.; Krcmar, H. (2017): A Competency Model for "Industrie 4.0" Employees, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 46-60

technologies and modified structures for communication and collaboration [3]. Processes will become interconnected and more complex. The technical, organizational and social spheres of work activities will overlap. The way we work will be one of the most affected changes in I4.0 [4]. I4.0 will not only affect technology and production, but the way we will work in all its dimensions [5].

This transformation of the work environment will change the job profiles and therefore requires employees to be outfitted with a wide range of competencies [1, 2, 6]. In I4.0, work profiles that require a higher education will gain increasing significance, while labour workforce will be mostly replaced by automated processes [1]. As a consequence, various practitioners and researchers agree that the competency development for students and employees applying for jobs that require higher education, is one of the key challenges to adapt I4.0 [3, 7-12]. To address this challenge, Erol, Jäger, Hold and Sihn [9] propose competencies derived from the literature by offering a scenario-based learning concept for students. acatech, Fraunhofer Institut für Materialfluss und Logistik and equeo GmbH [2] analyzed German companies by following a holistic approach and propose a set of competencies divided in two areas: competencies that the companies should master and competencies that the employees should adapt. Other authors also analyze working in I4.0 by specifying competencies that will become important [4, 7, 13-15].

In order to successfully get through the transformation towards I4.0, a clear definition of the competencies for I4.0 is needed [7, 8, 16]. Furthermore, a clear description of the relationship and connection between these competencies can provide the foundation for competency development in the future [2]. The best way to address this point would be a structured competency model, which addresses I4.0 competencies for graduates. I4.0 is accompanied by the enhancement of production machines, which requires adjusted competency profiles for engineers. IT assumes the role of programming these machines and designing adjusted IT architectures, which requires new competencies for IT professionals. These changes in production, the transformation of business processes as well as new ways of communication and collaboration will lead to adjusted or even new IT processes and structures, but also to a different way of managing people, which requires customized competency profiles for Information System (IS) professionals. Job profiles for engineering, IT, and IS employees need to be adjusted and include new competencies in order to cope with I4.0.

Therefore, we address this research gap by identifying competencies for I4.0. We focus on three areas that require higher education and will be of high relevance in I4.0: IS, IT, and Engineering. Our research addressed the following research question:

RQ: What competencies are critical for job positions that require higher education for effectively and efficiently performing in I4.0?

We offer a competency model with three different variants for these three areas by combining two research methods: a literature review and focus groups.

In the next section of this paper we describe the main concepts used throughout this research: “Competency”, “Competency Model” and “Industry 4.0”. Afterwards we explain the applied methodology by describing in detail each of the applied research methods namely the literature review and the focus groups. In the following section we present the result of the research by describing the results delivered by each of the methods and the delivered competency model. We conclude the paper with a discussion of our findings.

2 Background

2.1 Competencies

Many disciplines of research, such as Psychology, Education, Organizational Management, Human Resources or Information Systems have studied the concept of competencies. Various researchers provided different definitions over the years and caused a debate that is still ongoing [17]. The first definition of competencies was delivered by McClelland [18], who defined a competency as “a personal trait or set of habits that leads to more effective or superior job performance”. On later years further definitions can be found in research, e.g., Klemp [19] defined a competency as “an underlying characteristic of a person, which results in effective and/or superior performance on the job”. With regards to Spencer and Spencer [20], “competencies are skills and abilities; things you can do; acquired through work experience, life experience, study or training”. Bartram, Robertson and Callinan [21] state that competencies are “sets of behaviors that are instrumental in the delivery of desired results or outcomes”.

Research on competencies has mainly followed three approaches that were developed independently [17]. The behavioral approach, focuses on attributes which go beyond the cognitive ability, like self-awareness, self-regulation and social skills [18, 22]. This approach argues that competencies are fundamentally behavioral unlike personality or intelligence and can be taught through learning and development. The functional approach focuses on competencies as requirements for successfully fulfilling a task by restricting the term of competencies to the skills and know-how required for conducting a task [23, 24]. The holistic/multi-dimensional approach describes competencies as a collection of individual competencies required from an individual – and organizational competencies required on the organization level to achieve the desired results [25].

In this study we focus on the individual as a key factor in I4.0, by analyzing the broad spectrum of competencies for individuals not only on functional but also on behavioral level. We do not define a list of skills for fulfilling a certain task and also do not address organizational competencies. Moreover we want to offer an overview of the competencies that should be taught to individuals for successfully working in I4.0. Therefore we apply the behavioral-based approach since it offers the best fit for our purpose, by giving also the possibility to describe the relationship between competencies as constructs on the one hand, and psychological constructs such as motives and personality traits on the other [26].

For the purpose of this study we use the definition of Bartram, who defines competencies as: “sets of behaviors that are instrumental in the delivery of desired results or outcomes” [21]. In this sense “a competency is not the behavior or performance itself but the repertoire of capabilities, activities, processes and responses available that enable a range of work demands to be met more effectively by some people than by others” [27].

2.2 Competency Models

A competency model consists of desired competencies for a certain task and may also include a description of single competencies [28-30] as well as indicators to measure performance and outcome. This lists may include different detail levels and also describe relationships between the competencies.

Many competency models have been developed over the years. E.g., Erpenbeck and Rosenstiel [31] offer a model by separating the competencies into four categories: personal, social/interpersonal, action-related and domain-related competencies. Egeling and Nippa [32] use another classification by separating competencies in meta, domain, method and social competencies. Other authors offer competency models for leadership and management [22, 24, 33]. There are also competency models for certain tasks or job profiles [26, 32].

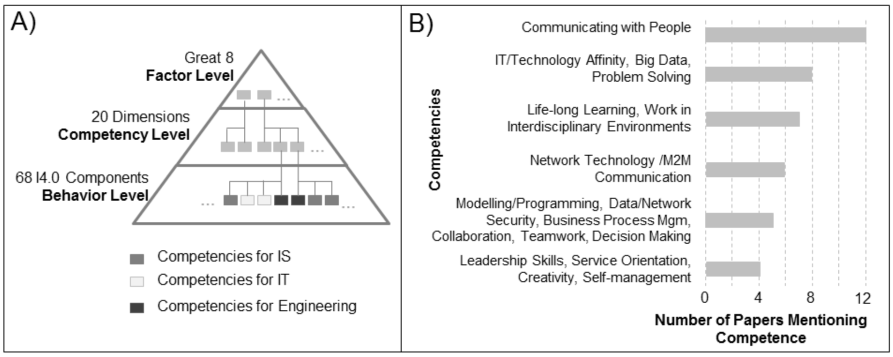


Fig 1. **A)** Industry 4.0 Competency Model Structure (Source: Own representation with regards to [37]); **B)** Most Mentioned Competencies in the Literature (Source: Own representation)

CEB Inc. [34][1] offers the SHL Universal Competency Framework (UCF)[2] [35] as a generic foundation for building competency models. This behavioral-based framework was derived by analyzing practitioners and academic approaches. It consists of three hierarchical levels, with the first level called “Great Eight”. It describes the eight core factors that underpin job performance. All competencies can be clustered in these eight groups of competencies, followed by 20 competency dimensions that divide these eight groups in further categories, which are separated into 112 component competencies. At this level all available competencies are described and each competency can be matched

1 CEB Inc. is a global best practice and insights technology company providing services to businesses worldwide [34].

2 SHL Universal Competency Framework (UCF) presents a state-of-the-art perspective on competencies and is used worldwide from well-known companies as e.g. Coca Cola [36]. It is offered by CEB Inc (see above) [35].

at one of the 112 elements (Fig. 1A). Moreover, it offers a deep level of detail. This framework offers a general perspective on competencies, within which competency models for concrete topics can be developed. For our model we adapted the competency framework (see Fig. 1A) by using the "Great Eight" as the first level and the 20 competency dimensions as the second level. We adapted the needed competencies of the behavior level as third level based on our results from the literature and focus groups. This way we based our model on a well-known framework from practice and research and adapt this framework for the I4.0 needs.

Choosing this existing framework offers many advantages. It offers a state of the art structure for competency modeling by not only listing the competencies but also showing the relationships between them. The framework is used both in research and practice, so our work makes a two-fold contribution. Since many companies apply it to build their competency profile, it offers the potential to compare our results with industry profiles in practice.

2.3 Industry 4.0

I4.0, also known as the fourth revolution, is one of the ten future projects of the High-Tech-Strategy 2020 action plan that was announced the first time by the Federal Government at the „Hannover-Messe" in 2011[1]. Its aim is a more efficient, flexible and individual production, achieved through decentralized controls of production and completely digitally controlled or even self-organized value chains [4], and where automation, real-time and sensor technologies play a crucial role [1]. "Plattform Industrie 4.0"[38] defines I4.0 as:

„[...] the fourth industrial revolution, the next stage in the organization and control of the entire value stream along the life cycle of a product. [...] based on increasingly individualized customer wishes and ranges from the idea, the order, development, production, and delivery to the end customer through to recycling and related services. [...] availability of all relevant information in real-time through the networking of all instances involved in value creation as well as the ability to derive the best possible value stream from data at all times. Connecting people, objects and systems leads to the creation of dynamic, self-organized, cross-organizational, real-time optimized value networks, which can be optimized according to a range of criteria such as costs, availability and consumption of resources."

It should be noted that the term I4.0 is widespread in German speaking countries. However, similar concepts and visions are often used under another term in the international context. For instance, I4.0 is known by the term „industrie du futur" in France, or the „Industrial Internet"[39] as well as further similar concepts as „Internet of Things", „Internet of Everything" „Smart Factory" or "Digital Transformation" in the international context [40, 41]. All this concepts include the use of automation, real-time, sensors and further modern technologies to transform business processes and therefore achieve a business value, however they slightly differ from one another in various aspects. For the purpose of this study we refer to I4.0 as a German concept with regards to the definition presented above.

3 Methodology

3.1 Literature Review

To define competencies for I4.0 we conducted a systematic literature review, which offers a rigorous view of research results [42]. We chose a concept-centric approach by following the recommendations of Webster and Watson [43]. The main objective of the literature review was to identify, classify and summarize competencies about I4.0 that were defined in the literature.

Following the guidelines of Webster and Watson [43] we searched by using the keywords: “Industrie 4.0”; “Industry 4.0”; “Digital Transformation”; “Internet of Things”; “IoT”; “Cyber Physical Systems”; “CPS” and combined each of them with the keywords: “competence”, “competency”; “skill”; “knowledge”; “attitude”; “ability”; “value”; “education”. Our goal was to conduct an exhaustive literature search and cover the state-of-the art literature about I4.0 competencies. The chosen databases were ACM Digital Library, IEEE, Springer and EbscoHost[3] because they cover publications from the IS, Economics, IT and Engineering discipline, as well as many Education outlets including conferences like EDUCON, REV, ICL, and Frontiers in Education that are often target outlets for publishing competency studies regarding actual topics like I4.0. The search included all articles that were published until August 2016. All the hits were first screened based on the title and abstract. In a second phase, the whole articles were screened. Additionally, a Google Scholar search was conducted in order to discover relevant articles from conferences and journals that were not included in the databases mentioned above. Here, the articles were sorted by relevance and the first 30 hits for each search string were screened. By following the recommendations of Webster and Watson [43], a backward and forward search was also conducted from the analyzed articles. Articles that did not include concrete competencies were excluded from our analysis. We had a total of 3363 hits in the database search, after the first screening 26 articles from the databases remained for further analyzes. Only articles where explicit competencies are mentioned were chosen. At the end a total of 17 that mention competencies for I4.0 or similar concepts such as IoT, were selected for further analysis. One of the articles was from the backward search.

Since the topic is new, only little research exists. However we conducted a literature review to summarize the state of the art before gathering any further data. The topic is of high practical relevance and broadly discussed in practitioners’ texts. Therefore by following the recommendations of Levy and Ellis [44], we also considered practical articles, white papers and reports that propose competencies for I4.0. These were determined through Google search and delivered a total of 10 practical articles included in our analysis. Finally, 27 articles including research and practitioners’ publications were considered and analyzed.

[3] The used EbscoHost Databases are: Business Source Premier, EconLit, Information Science & Technology Abstracts, Education Source, ERIC, Business Source Complete

From each article we extracted the mentioned competencies and built a concept matrix as proposed by Webster and Watson [43]. If the same competency was covered as a synonym in different papers e.g. “smart data” and “big data” we considered this as one competency and used the more popular term.

3.2 Focus Groups

For evaluating and extending the literature review, we conducted focus group interviews as recommended by Krueger and Casey [45]. A total of four focus groups with 18 - 25 participants each were conducted. The focus groups lasted 45 minutes on average. The participants in the focus groups were lecturers with previous experience in companies or various years of experience in university teaching and education in the areas of IT, IS, Economics and Engineering. This target group was addressed since lecturers have a general understanding of competencies and apply competency targeted teaching. Most of them also are involved in research and therefore are aware of I4.0, its relevance and the importance of building up competencies for the future employees. Three of the focus groups included lecturers from different countries in the EMEA region e.g. Germany, Austria, Netherlands and Egypt. They were conducted at the Technical University of Munich during training workshops for lecturers, who are interested in modern technologies that can be applied for teaching purposes, including topics of digital transformation, IoT and I4.0. The last focus group included professors and lecturers from Germany, Austria and Switzerland. It was conducted during a workshop at a German software company that aimed, among other topics, in discussing challenges and technologies that should be applied in today’s education. Using the setting of a workshop was helpful since each group had already known each other during the workshop and built a certain group dynamic that positively influenced the discussion. The participants also had time to discuss and think about I4.0 related topics during the workshops, so they were in the right mindset for the discussion and for building up ideas. Due to the given group dynamic and workshop setting we decided to limit the focus groups to the given lecturers and professors, since most of them had also practical insights from their previous jobs.

All focus groups have been moderated by the same person, a co-author of the paper. We used the same semi-structured guidelines in each of the focus groups to ensure that the findings are comparable. We applied the Critical Incident Technique [46, 47] for the focus group guidelines in order to derive the competencies for I4.0. The participants were presented with typical work scenarios and products of I4.0. Then questions were asked about the competencies that employees should bring in Engineering, IT and IS to efficiently work in this scenario.

The focus groups were recorded and transcribed. We coded the transcripts using the software MAXQA and combined an inductive with a deductive coding approach. This means we took the competencies from the literature as codes and started coding the transcripts. If a new competency was mentioned in a focus group that was not part of the codes, we used this as a new code meaning a new competency in our list. The coding was conducted twice from two different researchers. The codes were lastly compared and the differences were discussed until a common decision on the code was achieved.

4 Results

4.1 Results from Literature

Based on the literature, a total of 64 competencies could be derived. Most of them were behavioral ones and underline the importance of behavioral competencies for I4.0. It cannot be expected that a single person possesses all the mentioned competencies. So different combinations of the competencies represent specific job profiles for I4.0. The most mentioned competencies and their occurrence in the analyzed literature are presented in Fig. 1B.

Various authors underline that communication is one of the key competencies required from graduates [1, 2, 6, 7, 9, 13, 48-53]. Others go further by putting the communication competency in relation with other competencies like literacy [50] and technical communication [9, 50], intercultural competency [9, 13, 50], or presentation ability [54]. Social skills like collaboration [2, 7, 49, 54, 55], compromising [9], and negotiating [55] combined with emotional intelligence [55] will play a key role in I4.0 since they also play an important aspect in teamwork [7, 9, 48, 49, 51], project management [52, 53], and management ability [6], customer orientation [2, 13], maintaining customer relationships [2, 56], and creating business networks [2, 56]. Work and collaboration will become more complex, therefore I4.0 requires graduates with analyzing competencies like problem solving [2, 4, 6, 7, 9, 15, 49, 55], optimization [2, 4], analytical skills [9, 12, 57], and cognitive abilities [55]. To be able to coordinate these competencies, being able to manage complexity [2, 9] and abstraction ability [6, 9, 15] are crucial. Graduates in I4.0 should bring leading and deciding competencies like decision making [1, 2, 6, 55, 58], taking responsibility [6] and leadership skills [2, 6, 12, 55], which should be combined with a set of principles and values with competencies like respecting ethics [51], environmental awareness [52, 53], and awareness for ergonomics [48].

I4.0 will lead to a dynamic, international and interdisciplinary work environments, therefore competencies such as working in interdisciplinary environments [2, 4, 7, 12, 48, 51, 54], flexibility [9], adaptability [48, 49] as well as innovating [2, 14], creativity [7, 9, 14, 49], critical thinking [49], and change management [56] gain a new importance. For being able to always adapt the latest technologies and make the most out of them, graduates should apply life-long learning [4, 9, 12, 49, 50, 51, 54] and knowledge management [48, 49] while being focused on business strategy [3], always changing business models [3, 54] and entrepreneurship [49]. The work environment will become very demanding, so a graduate will need to find work-life balance [9] and needs to have the competency of self-management and organization [1, 6, 48, 49] as well as of planning and organizing work [13, 49, 57]. Nevertheless he should bring legislation [48, 52, 53] and safety awareness [51, 57] as well as individual responsibility [6].

Apart from all the behavioral competencies mentioned above, graduates must also bring domain related competencies as well as the ability to apply expertise and use technology. In this area all graduates need to bring IT and technology affinity [2, 4, 9, 12, 13, 48, 56, 57], economics knowledge [52, 53], and be able to extract business value from the use of social media [9, 56]. IS graduates should have knowledge in service orientation and product service offerings [2, 3, 56, 59], business process [2, 3, 9, 48, 54], and change management [56]. IT graduates should have knowledge of digital security, including data and network [2, 3, 51, 56, 59], and while working with engineers both groups should bring the competency of integrating heterogeneous technologies [51-53], knowledge about mobile technologies [56] and embedded systems and sensors [51], knowing network technology and M2M communication [2-4, 9, 54, 59] as well as possess knowledge of robotics and artificial intelligence [2, 12, 57]. On the other hand, IT and IS graduates should both bring modelling and programming knowledge [9, 12, 48, 58, 60], knowledge about cloud computing and cloud architectures [2, 56, 59], in-memory DB knowledge [56] and statistics [48]. For both groups, big data and data analysis and interpretation [2, 3, 9, 12, 48, 54, 56, 61] will be of big importance.

4.2 Focus Group Results

The most mentioned competency in the focus groups was big data/data analytics competency. "*I think it is about all different kind of data, also geo data but also video data, images, all ERP data, structured data and unstructured data like Facebook etc." "So to use anonymized big data and volume data and data traffic to predict macro business events rather than micro."* The participants see this as the next big thing and believe that for succeeding in I4.0, a combination of big data competency with sensors and mobile technology as well as predictive maintenance and machine learning will be very important.

The next most mentioned competency is process know-how and process management competency. De facto, processes are in the focus of I4.0, with automation playing an important role. The participants also underlined that business model understanding and entrepreneurship will play a special role in I4.0 since this will transform business models. The employees of tomorrow should be prepared to use the technological advances as an advantage and adapt in a fast changing world. "*The question is: Which potential does the digitalization bring and which new services can be offered based on that?*"

The participants also stressed that interdisciplinary competency will play a new role in I4.0. An engineer will have to collaborate with the IS and IT specialists in order to achieve results in the interconnected environment that we will face.

The domain or analytical oriented competencies like IT and technology affinity, network administration, data security cloud architectures, programming, in-memory DBs were also important in the discussion. *"Just to have the picture. You run through the world and Industry 4.0, you know there are so much sensors." "The more technical people they should know afterwards how to create systems."*

Lastly the participants also mentioned further behavioral competencies like customer orientation, decision making, communication, innovating, legal, ethics, and teamwork. *"I do not need to understand the whole technical background, but I need to be able to make decisions." "...we should offer group work, so that the participants learn to communicate and work in teams".*

Compared to the literature we could add four new competencies that were mentioned in the focus groups but have not been found in literature, as customer relationship management, IT architectures, machine learning, and predictive maintenance. Competencies in the dimension of leading, like leadership skills, or persuading and influencing like negotiating and emotional intelligence that were often mentioned and analyzed in the literature have not been mentioned at all during the focus groups. The mentioned competencies were generally more abstract, lacking high detail compared to the literature.

4.3 Developing the Model

The SHL UCF[2] developed by CEB Inc.[1] is based on different competency approaches from research and practice [35]. It offers a behavioral approach for competency modeling by focusing on the individual and considering competencies of behavioral nature, meaning an individual can learn and adopt them unlike, e.g., personality. As a framework it offers a structure and overview of the competencies, by fitting them into descriptive categories [62]. This framework can be used to develop competency models, which represent a descriptive and simplified view of the competencies as a specific phenomenon to be analyzed [62]. The SHL UCF is widely used in practice and many companies use it to describe their competency models for specific job positions [26].

As described earlier, the SHL UCF is composed of three hierarchical levels: the "Great Eight", the competency dimensions and the competency components. We kept the structure and the relationship between the elements and adapted the third level competencies based on the results of our research. The framework delivers 112 single competencies. For I4.0, we considered 68 competencies (Fig. 1A) as relevant. Based on the results of the literature review and focus groups we expanded some of the single competencies or adapted their formulation to serve our purpose. For each of the competencies we did a clustering whether it is relevant for IS, IT or Engineering graduates or whether it can be considered as an interdisciplinary competence for two of the mentioned areas or all three of them. The process of clustering was conducted separately from two researchers and then compared to each other. In case of disagreement the clustering was discussed until achieving a consensus.

The results show that most of the behavioral competencies should be adapted by all three groups of graduates. These competencies are marked in light grey color eg. *Decision Making* or *Teamwork*. It means that the employees of the future, independently from their position should bring a high level of behavioral competencies to successfully work in I4.0. Only competencies under the dimension "Applying Expertize and Technology" have three variants. This dimension of competencies represents domain knowledge, therefore depending on the domain the employees should bring different competencies. Some competencies in this dimension are also categorized to two or more groups of graduates. E.g., *Predictive Maintenance* will be a competency for IT as well as for Engineering graduates, whereas *Big Data* will be a needed competency not only for IS but also for IT graduates. Economics graduates, who follow a technical oriented carrier path, will adapt similar competencies as the IS

<table>
<tr><th>Big Eight</th><th>Competency Dimensions</th><th colspan="3">Competencies</th></tr>
<tr><th></th><th></th><th><u>Information Systems (IS)</u></th><th><u>Computer Science</u></th><th><u>Engineering</u></th></tr>
<tr><td rowspan="2">Leading & Deciding</td><td>Deciding and Initiating Action</td><td colspan="3">• Decision Making
• Taking Responsibility</td></tr>
<tr><td>Leading and Supervising</td><td colspan="3">• Leadership Skills</td></tr>
<tr><td rowspan="2">Supporting and Cooperating</td><td>Working with People</td><td colspan="3">• Teamwork
• Collaborating with Others
• Communicating with People</td></tr>
<tr><td>Adhering to Principles and Values</td><td colspan="3">• Respecting Ethics
• Environmental Awareness
• Awareness for Ergonomics</td></tr>
<tr><td rowspan="3">Interacting and Presenting</td><td>Relating and Networking</td><td colspan="3">• Compromising
• Creating Business Networks
• Maintaining Customer Relationships</td></tr>
<tr><td>Persuading and Influencing</td><td colspan="3">• Negotiating
• Emotional Intelligence</td></tr>
<tr><td>Presenting and Communicating Information</td><td colspan="3">• Presentation and Communication Ability</td></tr>
<tr><td rowspan="6">Analyzing and Interpreting</td><td>Writing and Reporting</td><td colspan="3">• Targeted/Technical Communication
• Literacy</td></tr>
<tr><td rowspan="4">Applying Expertise and Technology</td><td colspan="3">• IT and Technology Affinity
• Economics
• Extract Business Value from Social Media</td></tr>
<tr><td>• Service Orientation/ Product Service Offerings
• Business Process Management
• Business Change Management
• Understand and Coordinate Workflows</td><td>• Network Security
• IT Architectures
• Machine Learning</td><td></td></tr>
<tr><td></td><td colspan="2">• System Development
• Integrating Heterogeneous Technologies
• Mobile Technologies
• Sensors/Embedded Systems
• Network Technology /M2M Communication
• Robotics/Artificial Intelligence
• Predictive Maintenance</td></tr>
<tr><td colspan="2">• Modelling and Programming
• Big Data/Data Analysis and Interpretation
• Cloud Computing /Architectures
• In-Memory DBs
• Statistics
• Data Security</td><td></td></tr>
<tr><td>Analyzing</td><td colspan="3">• Problem Solving
• Optimization
• Analytical Skills
• Cognitive Ability</td></tr>
<tr><td rowspan="3">Creating and Conceptualizing</td><td>Learning and Researching</td><td colspan="3">• Life-long Learning
• Knowledge Management</td></tr>
<tr><td>Creating and Innovating</td><td colspan="3">• Innovating
• Creativity
• Critical Thinking
• Change Management</td></tr>
<tr><td>Formulating Strategies and Concepts</td><td colspan="3">• Business Strategy
• Abstraction Ability
• Managing Complexity</td></tr>
<tr><td rowspan="3">Organizing and Executing</td><td>Planning and Organizing</td><td colspan="3">• Project Management
• Planning and Organizing Work
• Management Ability</td></tr>
<tr><td>Delivering Results and Meeting Customer Expectations</td><td colspan="3">• Customer Orientation
• Customer Relationship Management</td></tr>
<tr><td>Following Instructions and Procedures</td><td colspan="3">• Legislation Awareness
• Safety Awareness
• Individual Responsibility</td></tr>
</table>

Big Eight	Competency Dimensions	Competencies		
		IS/Economics	IT/Computer Science	Engineering
Adapting and Coping	*Adapting and Responding to Change*	• Work in Interdisciplinary Environments • Intercultural Competency • Flexibility • Adaptability and Ability to Change Mind-set		
	Persuading and Influencing	• Work-Life Balance		
Enterprising and Performing	*Achieving Personal Work Goals and Objectives*	• Self-management and -organization		
	Entrepreneurial and Commercial Thinking	• Business Model Understanding • Entrepreneurship		

Fig 2: "Industrie 4.0" Competency Model (Source: Own representation with regards to [35])

graduates, since these disciplines have similarities. This shows again that the work in I4.0 will be interconnected. Therefore, competencies such as interdisciplinary working, collaboration, communication or teamwork will have a special role.

Our model is presented in Fig. 2. For each of the employee groups you can follow the path and gather all the competencies that should be fulfilled by this group. It cannot be expected that one employee of a certain group masters all the competencies. Therefore, a combination of the competencies, depending on the position will deliver different job profiles for I4.0. E.g. a competency profile for a data scientist responsible for extracting, modeling and visualizing the data produced by a certain sensor in I4.0 can be defined by extracting concrete competencies from the area of IS such as *Taking Responsibility*, *Big Data Analytics and Interpretation*, *Analytical Skills*, *Cognitive Ability*, *Creativity*, and *Critical Thinking*. By following this schema different profiles for different jobs could be defined.

5 Discussion and Limitations

In our research we used the SHL UCF [35] to develop an I4.0 competency model. Most of the defined competencies are not new, however the presented specific combination of competencies for I4.0 is new and makes a contribution to research. Overall, our research emphasizes the importance of employee competencies to successfully get through the transformation towards I4.0. The results of the literature review and the focus group discussions delivered mostly behavioral competencies and only a small part of the competencies represented domain related knowledge. This is also a new aspect with regards to competency building and underlines the changes that I4.0 will bring to the way we work. Job vacancies in today's economy often focus on a list of domain knowledge and comprise only some very generic behavioral competencies, like teamwork or independently working. The same situation is presented if we analyze university lectures and curricula. Their focus often is on teaching the students' domain knowledge. The training of further competencies still often is limited to teamwork situations or presentations to be held. These examples show that domain knowledge are the focus of today's economy, while I4.0 will turn around the work environment. Behavioral competencies will be the most important competencies that employees should bring. Therefore, research should focus on analyzing how the competency profile of today's employees as well as of students could be adapted for I4.0. This could include the definition of requirements for curricula and training programs for I4.0. Conceptualizing and defining learning strategies and curricula for I4.0 might be another interesting topic for research.

If we consider today's economy and disciplines, there is a clear separation between the competencies that employees from different disciplines should bring. If, for instance, we mention IT, everyone recalls a certain job profile and competency set in their mind that is completely different from the profile that one would recall if we mentioned IS or Engineering. Our study reveals that in the future the competency sets that different disciplines should bring will be very similar and will differentiate only in some aspects of domain knowledge. This would be a further point where research could offer teaching methods for interdisciplinary teaching.

Our literature review also showed that research on I4.0 competencies is rather scarce. The analyzed works mostly underline that the work environment will change, however no concrete vision or competency models were proposed. With our work we make an initial contribution that could be further expanded for other professions that require higher education.

This work has practical implications as well. The proposed competency model could be used in practice by companies and universities. Companies could use the model to define job profiles for I4.0 vacancies. It cannot be expected that one employee will bring all the competencies included in the model, however by combining some of them depending on the position, different profiles can be described. The results can also be used in competency-based curricula designing.

Although there are limitations to our study, we believe that it can serve as a foundation for further research. Our analysis was based on a literature review as well as four focus groups with academic staff. The literature search was limited to the databases, where access by our university, is provided. To complete the results further literature, especially conference proceedings as well as empirical data e.g. focus groups or expert interviews from practitioners would be helpful. We acknowledge that further research in the area of I4.0 competencies is required to study further aspects of competencies as well as define how the model could be applied in practice. Especially a definition of a competency profile for a certain job description, e.g. which competencies of the model should a programmer bring, could be a further interesting point for research and practice.

References

1. Kagermann, H., Wahlster, W., Helbig, J.: Recommendations for implementing the strategic initiative Industrie 4.0. Report, Industry 4.0 Working Group (2013)
2. acatech, Fraunhofer Institut für Materialfluss und Logistik, equeo GmbH: Kompetenzentwicklungsstudie Industrie 4.0. Report (2016)
3. Zinn, B.: Conditional variables of 'Ausbildung 4.0'. JOTED 3, 1-9 (2015)
4. Gebhardt, J., Grimm, A., Neugebauer, L.M.: Developments 4.0 Prospects on future requirements and impacts on work and vocational education JOTED 3, 117-133 (2015)

5. Bundesministerium für Arbeit und Soziales (BAS).: Arbeiten 4.0. Report, BAS (2015)
6. Smit,J., Kreutzer,S., Möller,C.,Carlberg,M.:Industry4.0.Report. European Parliament (2016)
7. Richter, A., Heinrich, P., Stocker, A., Unzeitig, W.: Der Mensch im Mittelpunkt der Fabrik von morgen. HMD 52, 690-712 (2015)
8. Jaschke, S.: Mobile Learning Applications for Technical Vocational and Engineering Education. In: Int. Conf. on Interactive Collaborative Learning, pp. 603-608. Dubai (2014)
9. Erol, S., Jäger, A., Hold, P., Sihn, W.: Tangible Industry 4.0: a scenario-based approach to learning for the future of production. In: Conf. on Lear. Fact. Gjøvik, pp. 1-6. Norway (2016)
10. McKinsey&Company: Industry 4.0 - How to navigate digitization of the manufacturing sector. Report, McKinsey&Company (2015)
11. Deloitte: Making in an Industry 4.0 World. Report, Deloitte (2015)
12. The Boston Consulting Group: Man and Machine in Industry 4.0. Report, BCG (2015)
13. Guo, Q.: Learning in a Mixed Reality System in the Context of ‚Industrie 4.0'.JOTED 3, 92-115 (2015)
14. Stocker, A., Brandl, P., Michalczuk, R., Rosenberger, M.: Mensch-zentrierte IKT-Lösungen in einer Smart Factory. Elektrotechnik und Informationstechnik 2014, 207-211 (2014)
15. Windelband, L.: Zukunft der Facharbeit im Zeitalter „Industrie 4.0". JOTED 2, 138-160 (2014)
16. Richert, A., Shehadeh, M., Plumanns;, L., Groß;, K., Schuste, K., Sabina Jeschke: Educating Engineers for Industry 4.0. Global Eng. Education Conference, Abu Dhabi (2016)
17. Deist, F.D.L., Winterton, J.: What Is Competence? Human Res. Dev. Int. 8, 27-46 (2005)
18. McClelland, D.: Testing for Competence Rather Than for "Intelligence". America Psychologyst 28, 1-28 (1973)
19. Klemp, G.: The assessment of occupational competence. Report. Nat. Inst. of Edu. (1980)
20. Spencer, L., Spencer, S.: Competence at Work: Model for Superior Performance. John Wiley & Sons, New York (1993)
21. Bartram, D., Robertson, I.T., Callinan, M.: Introduction. A framework for examining organizational effectiveness. In: Robertson, I.T., Callninan, M., Bartram, D. (eds.) Organizational Effectiveness. The Role of Psychology, pp. 1-10. John Wiley & Sons, Baffins Lane, Chicheser, UK (2002)
22. Boyatzis, R.E.: The Competent Manager. Wiley, New York (1982)
23. Frank, E.: The UK's Management Charter Initiative: the first three years. Journal of European Industrial Training 17, 9-11 (1991)
24. Miller, L.: Managerial competences. Industrial and Commercial Training 23, 11-15 (1991)
25. Straka, G.A.: Measurement and evaluation of competence. Report, Cedefop (2004)
26. Kleindauer, R., Berkovich, M., Gelvin, R., Leimeister, J.M., Krcmar, H.: Towards a competency model for requirements analysts 395 1.2. Inf.Sys.Jor.2012, 475-503 (2012)
27. Kurz, R., Bartram, D.:Competency and individual performance.In: Robertson, I.T., Callninan, M., Bartram, D. (eds.) Organizational Effectiveness. pp. 227–255. Wiley, UK (2002)
28. Mirabile, R.L.: Everything you wanted to know about competency modeling. Training and Development 73-77 (1997)
29. Lucia, A.D., Lepsinger, R.: The art and science of competency models: Pinpointing critical success factors in organizations. Jossey- Bass/Pfeiffer, San Francisco (1999)
30. Markus, L., Cooper-Thomas, D., Allpress, N.: Confounded by Competencies? New Zealand Journal of Psychology 34, 117-126 (2005)
31. Erpenbeck, J., Rosenstiel, L.: Handbuch Kompetenzmessung. Schäffer Poeschel (2007)
32. Egeling, A., Nippa, M.: Kompetenzbedarfe im Kontext hybrider Wertschöpfung. In: Reichwald, R., Krcmar, H., Nippa, M. (eds.) Hybride Wertschöpfung. Eul, J, Germany (2009)
33. Basellier, G., Reich, B.H., Benbasat, I.: Information Technology Competence of Business Managers: A Definition and Research Model. Jour of Mgm. Inf. Systems 17, 159-182 (2001)

34. CEB Inc., http://www.shl.com (Accessed: 22.07.2016)
35. Bartram, D.: The great eight competencies: a criterion-centric approach to validation. Journal of Applied Psychology 90, 1185–1203 (2005)
36. Bartram, D.: The SHL Universal Competency Framework. Report, SHL Group (2011)
37. Iliescu, D.: Competence assesment practices in SHL, SHL Group (2012)
38. bitcom, VDMA, ZWEI: Implementation Strategy Industrie 4.0. Report, bitcom, VDMA, ZWEI (2016)
39. Grangel-González, I., Halilaj, L., Coskun, G., Auer, S., Collarana, D., Hoffmeister, M.: Towards a Semantic Administrative Shell for Industry 4.0 Components. arXiv (2016)
40. Roth, A.: Industrie 4.0 – Hype oder Revolution? In: Roth, A. (ed.) Einführung und Umsetzung von Industrie 4.0, pp. 1-15. Gabler Verlag, Berlin
41. Obermaier, R.: Industrie 4.0 als unternehmerische Gestaltungsaufgabe. In: Obermaier, R. (ed.) Industrie 4.0 als unternehmerische Gestaltungsaufgabe, pp. 3-34. Gabler, Berlin (2016)
42. vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the Giant. In: ECIS, pp. 2206 - 2217. Verona (2009)
43. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26, xiii - xxiii (2002)
44. Levy, Y., Ellis, T.J.: A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research. Informing Science 9, 181 - 212 (2006)
45. Krueger, R.A., Casey, M.A.: Focus groups. Thousand Oaks: Sage (1994)
46. Flanagan, J.C.: The critical incident technique. Psychological Bulletin 51, 327–358 (1954)
47. Koch, A., Stroebel, A., Kici, G., Wesrhoff, K.: Quality of the Critical Incident Technique in practice. Psychology Science Quarterly 2009, 3-15 (2009)
48. VDI & ASME: Industry 4.0. Report, VDI & ASME (2015)
49. Kiesel, M., Wolpers, M.: Educational challenges for employees in project-based Industry 4.0 scenarios. i-KNOW, Graz, Austria (2015)
50. Xia, S.: Training Programs for Excellent Engineers with Engineering of Internet of Thing. In: Int. Symposium on IT in Medicine and Education, pp 610-615. Cuangzhou (2011)
51. Grega, W., Kornecki, A.J.: Real-Time Cyber-Physical Systems-Transatlantic Engineering Curricula Framework. In: Conf. on Comp. Sc. and Inf. Sys., pp. 755-762. Gdansk (2015)
52. Grimheden, M.E., Törgren, M.: Towards curricula for Cyber-Physical Systems. In: Workshop on Embedded and Cyber-Physical Systems Education, New Delhi, India (2014)
53. Maenpaa, H., Tarkoma, S., Varjonen, S., Vihavainen, A.: Blending Problem and Project Based Learning in IoT Education. In: Tech.Sym. on Com. Sc. Ed., pp. 398-403. USA (2015)
54. Roland Berger Strategy Consultants: Industry 4.0. Report, Roland Berger (2014)
55. Gray, A. https://www.weforum.org (Accessed: 22.07.2014)
56. Hoberg, P., Krcmar; H., Oswald, G., Welz, B.: Skills for Digital Transformation. TUM (2015)
57. Hartmann, E.A., Bovenschulte, M.: Skills Needs Analysis for “Industry 4.0” Based on Roadmaps for Smart Systems. In: SKOLKOVO (ed.) (2013): Using Technology Foresights for Identifying Future Skills Needs. pp. 24-37. Moscow (2013)
58. Kortuem, G., Arosha K., Smith, N., Richards, M. Petre, M.: Educating the Internet-of-Things generation. Computer 46, 53-61 (2013)
59. Chunzhi, W., Hui, X., Xia, M.: Construction of Hardware Curriculum Group for Transition from Network to IoT Engineering Major. Int. Conf.on Com. Sc. & Ed., pp. 1575-1579. Australia (2012)
60. Chin, J., Callaghan, V.: Educational Living Labs. Int. Conf. on Intelligent Environments, pp. 92-99. Athens (2013)
61. Capgemini Consulting: Industry 4.0. Report, Capgemini Consulting (2015)
62. Frankfort-Nachmias, C., Nachmias, D., Dewaard, J.: Res.Meth. in theSoc.Sc.Palgrave (2014)

Assessing Tools for Coordinating Quality of Master Data in Inter-organizational Product Information Sharing

Thomas Schäffer[1] and Dirk Stelzer[2]

[1]Hochschule Heilbronn,
Studiengang Wirtschaftsinformatik,
Heilbronn, Germany
Thomas.Schaeffer@hs-heilbronn.de

[2] Technische Universität Ilmenau,
Fachgebiet Informations- und Wissensmanagement,
Ilmenau, Germany
Dirk.Stelzer@tu-ilmenau.de

Abstract. Product information sharing, i.e., inter-organizational transfer of master data relating to products, is a problematic, error-prone, labor-intensive, and costly process in many companies. This paper presents findings of a focus group interview and case studies at three wholesale trading companies that share product information with hundreds of suppliers. We identify and assess coordination mechanisms and tools used to facilitate product information sharing. Spreadsheet files, e-mail messages, telephone calls, and personal meetings are predominant coordination tools. EDI connections, product identification and classification standards, online product catalogs, and data pools are not widely adopted in the trading organizations covered by our study. Reasons for the low adoption rate are that employees responsible for master data quality are either unaware of these resources or that they are convinced that the tools are too cost-intensive or not flexible enough.

Keywords: Product Information Sharing, Data Quality, Master Data, Coordination Mechanism, Coordination Tool

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Schäffer, T.; Stelzer, D. (2017): Assessing Tools for Coordinating Quality of Master Data in Inter-organizational Product Information Sharing, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 61-75

1 Introduction

Interoperability is a major prerequisite for the digitization of value chains. Legner and Wende [1] define business interoperability as „the organisational and operational ability of an enterprise to cooperate with its business partners and to efficiently establish, conduct and develop IT-supported business relationships with the objective to create value." [1]. Legner and Wende emphasize, that "being 'interoperable' refers to being able to share information between business partners, understand and process exchanged data, [and] seamlessly integrate it into internal ICT systems." [1] Achieving and maintaining an adequate level of quality of master data is a crucial precondition for business interoperability [2–5].

Intra-organizational issues of ensuring and maintaining quality of master data have gained extensive attention in the literature [6, 7]. However, achieving an adequate level of master data quality in inter-organizational business processes – particularly in product information sharing – has had much less attention [2, 5, 8, 9].

Product information can be defined as a set of data, e.g., identification number, weight, size, etc., that represents the product [2, 5, 9]. Product information sharing denotes the inter-organizational transfer of product information, a concept labeled "product information supply chain" by Legner and Schemm [8]. The term master data refers to critical business objects of an organization [3]. It describes products, suppliers, customers, employees, and similar objects that rarely undergo changes [10].

Data quality denotes the extent to which data meet specified requirements. Data quality is "a measure of the adequacy of the data for specific requirements [...] Data quality is a multidimensional, contextual concept, as it cannot be described with a single feature, but on the basis of different data quality dimensions and measures" [7]. A key quality characteristic is "fitness for use" [2, 6]. In a study of product information sharing de Corbière [2] identified four major data quality dimensions, namely, accuracy, completeness, timeliness and security.

The objective of our study is to identify and to assess mechanisms and tools for coordinating quality of master data in inter-organizational product information sharing. We aim to answer two research questions:

- R1: What mechanisms and tools do trade companies apply to coordinate master data quality in product information sharing with suppliers?
- R32: What strengths and weaknesses of coordination mechanisms and tools do experts perceive?

This paper is organized as follows. The next section presents findings of our literature survey. Section 3 explains the research methodology. In section 4 we describe findings of three case studies: processes, mechanisms and tools for coordinating quality of product information exchanged between wholesale trading companies and their suppliers. Section 5 presents and discusses key findings of our study. Finally, we conclude by summarizing implications for practitioners and by pointing out research opportunities.

2 Prior Research

We conducted a literature survey to identify prior research into arranging for quality of master data in inter-organizational product information sharing and to describe the state of the art in this field. In contrast to a literature review, a literature survey gives an overview of the relevant literature but does not provide a detailed analysis. Due to space limitations we cannot describe the method and our findings in full detail. Therefore, we shall limit the description to some key features here. We followed the guidelines provided by Webster and Watson [11] and vom Brocke et al. [12] to identify relevant publications. As a first step, we examined IS journals and IS conference proceedings using the AIS Electronic Library, ScienceDirect, Google Scholar, and SpringerLink. We conducted electronic searches in titles and abstracts on the following search term: [("data" AND "quality") AND (("inter-organizational" OR "business-to-business" OR "supply chain") OR ("product" AND "information" AND "sharing") OR ("data exchange"))]. These searches identified a total of 175 publications. After analyzing each article's abstract, keywords, or the full article when necessary, we excluded 144 articles that were duplicates or did not appear to be concerned with or relevant to our research focus. As a third step, we performed a forward and backward search in relevant articles to identify further sources that had not been identified by the previous step. A total of 39 publications was read in full and coded. We excluded all publications that only stated the keywords mentioned in the search term without elaborating on these concepts. Out of the 39 coded articles, 12 include passages of interest. The following statements present key findings of our survey.

- Data quality is a critical success factor for efficient cross-company collaboration [5, 13–16]. This is particularly the case for product information sharing in supply chains [2, 8, 15].
- Various studies have found that manufacturers, wholesale trading companies, and retailers are concerned about the quality of product information shared between trading partners [2, 8].
- Several surveys have shown that poor quality of product data exchanged between cooperation partners may lead to substantial cost increases or loss of sales [8, 13].
- The attempt to improve product data quality is a key driving force for using electronic exchanges of product information, e.g., the Global Data Synchronization Network (GDSN) data pool [2, 9]. Nakatani et al. [13] define a data pool as a repository that supports trading partners in obtaining, maintaining, and exchanging information on trade items in a standard format through electronic means.
- Using inter-organizational information systems for synchronizing data, e.g., data pools, electronic catalogs or marketplaces, often does not lead to the promised benefits. Trading partners also need to take steps to improve data quality internally [2, 5, 13]. Several companies are still in the process of setting up appropriate coordination mechanisms for managing their internal product information supply chain [8]. Dalmolen et al. [5], for example, found that IT personnel in large(r) organizations often have little awareness of the product information sharing

process in their own organizations. Likewise, business units lack sufficient IT knowledge to initiate potential improvements.

- Several authors [5, 9, 17, 18] found that standards for product identification and classification, e.g., the Global Trade Item Number (GTIN) and the United Nations Standard Product and Services Code (UNSPSC), and data pools, e.g., the GDSN data pool, have not gained wide acceptance in many industries because they do not meet key requirements of trading companies.
- The exchange of product information between trading partners is often supported by telephone calls, fax or e-mail exchanges. These processes are problematic, error-prone, labor-intensive, and costly [2, 8, 9, 13, 19]. Legner and Schemm [8] cite various studies that report significant direct labor costs due to the manual transfer of product information and its administrative processing as well as indirect effects of poor data quality on the supply and demand chain. Falge et al. [16] suggest that trading partners should establish a process for agreeing on common standards and systems and for defining service level agreements for data quality.
- Several scholars [2, 5, 14, 16, 19] encourage further research into data quality in the inter-organizational exchange of product information. Legner and Schemm [8] suggest qualitative and quantitative studies into sharing of product information between suppliers and retailers of goods. They also ask for more extensive investigations of interdependencies and coordination mechanisms in the inter-organizational product information supply chain.
- In our literature survey we have identified only few studies [8] that have attempted to assess mechanisms and tools for coordinating quality of master data in inter-organizational product information sharing. Most of these studies focus on only one or on a limited set of coordination mechanisms and tools. To the best of our knowledge, no author has yet attempted to assess a broad range of tools for coordinating quality of master data in inter-organizational product information sharing.
- Most of the papers included in our survey do not explicitly build on a theory. In the rare cases, when a theory for analyzing master data quality management in inter-organizational product information sharing is used [8, 14], the authors build on coordination theory.

3 Research Methodology

Our research combines a focus group interview and case study-based research. A focus group interview is an advanced form of an expert interview. Researchers interview a group of experts and document the findings for later evaluation [20–22]. Focus group interviews are well suited for acquiring new insights and ideas, and for structuring research questions into manageable chunks. In section 2 we have shown that arranging for adequate quality of master data in inter-organizational product information sharing is a research domain that has gained only little attention from scholars. Focus group interviews are especially suited for obtaining several perspectives about a complex topic. They help to structure the research domain and to identify research questions for further investigation.

In March 2016 we conducted a focus group interview with 14 master data experts from 12 enterprises. Five are trading companies, four manufacturers, and three IT service providers. All companies are dependent on high quality master data provided by hundreds of business partners. Each company has established a team or a project with the objective to improve inter-organizational product information sharing. The focus group deals with the exchange of product master data between manufacturers and retailers. The interview helped us to identify practically relevant research fields. The interview revealed that the process of setting up products in trading companies' product databases frequently leads to severe problems due to inadequate quality of master data provided by suppliers. Therefore, ensuring an appropriate level of quality of master data is of key importance for maintaining efficient supply chains among suppliers and wholesale trading companies. This finding is consistent with the research results described in section 2 of this paper. For this reason, we focused on this field in the subsequent case study research.

Case study research examines complex and ill-structured phenomena in a practical environment [23, 24]. Our research follows a single case study approach in order to extend our knowledge of the inter-organizational exchange of product master data. This approach was chosen due to little empirical research in the domain of inter-organizational master data quality as mentioned in section 2. The case study research was structured as follows: Similar to the research approach by Legner and Schemm [8], the selection of case study organizations was mainly driven by purposeful sampling, availability of multiple sources of information, and the willingness of experts to cooperate. We decided to focus on wholesale companies as these entities act as intermediaries between suppliers and retailers of goods. Thus, wholesale companies are particularly dependent on the quality of master data relating to trade items.

We conducted semi-structured expert interviews and analyzed documents and IT artifacts provided by three case study organizations. Interview guidelines were created based on the results of a pretest with several business experts. The guidelines comprise 22 questions. The interview questions were sent by e-mail to the experts one week before interviews took place. The interviews were held between June and August 2016. We talked to a total of 10 experts in three retail companies: master data experts, procurement staff and IT personnel. The interviews lasted between 4 and 6.5 hours. All interviews were recorded. Immediately after completing the interviews, the experts' answers were documented and coded. We then analyzed all findings following the recommendations of Mayring and Fenzl [25]. Subsequently, the interview documents were provided to the experts asking them to verify the findings and to resolve potential misunderstandings. The experts also provided us with internal documents, master data descriptions and spreadsheet files used to transfer product information among trading partners. These artifacts helped us to better understand specific details, processes, mechanisms, and tools for the inter-organizational coordination of master data quality in product information sharing.

4 Processes, Mechanisms and Tools for Coordinating Master Data Quality in Product Information Sharing

In the next paragraphs we describe processes, mechanisms and tools for coordinating quality of product information exchanged between wholesale trading organizations and their suppliers.

4.1 Case 1: Wholesale Trading Company in the Retail Market

The first study was carried out in a wholesale trading company in the retail market, hereafter referred to as C1. This company employs more than 5000 people and generates an annual turnover exceeding one billion Euros. C1 procures goods from more than 2000 suppliers and sells products to about 2000 customers worldwide, mostly retail stores. C1's product database consists of more than 300,000 items. Product master data sets include 109 attributes. About 120,000 master data sets are modified per year.

The organizational unit responsible for managing product master data, labeled MDM for short, is part of C1's fulfillment department. MDM consists of 18 staff members who are in charge of master data administration related to fulfillment tasks. At C1, fulfillment is defined as order handling process spanning from responding to sales inquiries over procurement to delivery of products to the retail sector.

The process of coordinating and agreeing on an appropriate level of master data quality with suppliers consists of several steps: (1) MDM creates an Excel file that serves as a container for product master data that C1 requests from suppliers. The file specifies about 100 attributes for each product. (2) C1's procurement department sends the file to suppliers by e-mail requesting to complete 20 to 30 attributes for each product. The exact number of required attributes depends on the product category. (3) 45 percent of all suppliers fill in the relevant data and return the completed file via e-mail. Approx. 55 percent of the suppliers do not fill in the Excel file. They provide the data in other formats, e.g., CSV or XML files, or on paper. Suppliers that do not provide data in Excel files cause substantial overhead for data import and/or data entry at C1. A limited set of suppliers run a customer portal. They provide master data for wholesale customers via the portal. In these cases, C1's MDM staff import data from the portal into the Excel file. (4) When master data provided by a supplier is received, C1's procurement department checks the data entries. About half of all files provided by suppliers are appropriate and correct. The other half is incomplete or defective. Inappropriate files lead to further inquiries via telephone or e-mail when C1's employees attempt to adjust and to complete product data. (5) Frequently, the Excel file needs to be returned to suppliers requesting to rectify master data sets. (6) In the next step, procurement staff members complement the data sets with about 30 further attributes. Characteristic examples are attributes that identify or categorize products, i.e., product identification numbers. (7) The Excel file is then forwarded to MDM. This unit carries out more quality checks. (8) In case of incomplete or false data entries MDM consults the procurement department or the supplier. During this step, the Excel file often has to be returned to the sender again with the request to complete or to amend master data sets. (9) When MDM employees consider the file appropriate they add another set of approx. 10 attributes, e.g., country-specific information needed for exporting products to international markets. (10) MDM than creates an upload file and loads the data into the database of the ERP system.

Figure 1 illustrates the process of inter-organizational product information sharing in UML notation as an activity chart.

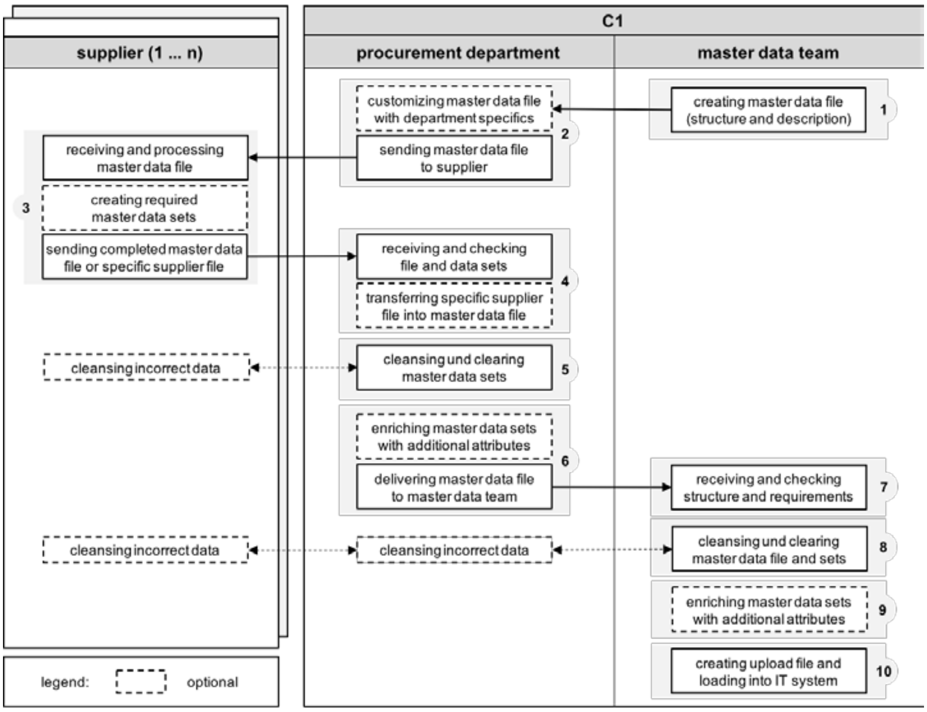


Figure 1: Overview of the process of product information sharing

4.2 Case 2: Wholesale Trading Company in the Automotive Aftermarket

The second case study was conducted in a wholesale trading company in the automotive aftermarket, hereafter called C2. The company has more than 5000 employees and generates approx. 1.6 billion Euros turnover per year. C2 purchases parts from more than 2000 suppliers and sells products to approx. 30,000 customers, mostly garages in five European countries. The stock of articles comprises 1.5 million items. Product master data sets consist of 92 attributes. 50,000 to 120,000 master data sets need to be modified per day.

The organizational unit responsible for managing master data is part of the procurement department and employs 8 people. They are in charge of master data administration related to all procurement tasks. However, master data team members usually do not directly communicate with suppliers. This is one of the duties of procurement staff.

Basically, C2's process of coordinating master data quality is similar to C1's process. In order to avoid repetitions, we refrain from describing the entire process again. We only mention specific features of C2's process in this section.

- C2's framework contracts with suppliers comprise a clause that mutually obligates trading partners to supply master data in adequate quality. However, this clause does not specify any details.
- C2's Excel file sent to suppliers consists of 92 attributes. 20 attributes are mandatory fields that must be filled in by suppliers.
- C2 utilizes the TecDoc catalog [26], a standardized electronic parts catalog that contains data sets for more than 5,400,000 products in the automotive aftermarket. Missing, incomplete or inaccurate data sets are complemented by C2's staff with data obtained from this catalog.
- Unlike C1, C2 utilizes several automatic routines to check quality of master data provided by suppliers before uploading master data sets into the product data base.
- Similar to the process at C1, C2's procurement staff members communicate with suppliers, request product master data and provide suppliers with Excel files to fill in the master data. Procurement employees frequently modify spreadsheet files sent to suppliers in order to better adapt to the specific requirements. This leads to an ambivalent situation: On the one hand the Excel files are tailored to the specific requirements of procurement units and to certain product categories. On the other hand, procurement staff members do not communicate these modifications to personnel in the master data unit. This leads to inconsistencies in master data sets, to frequent check backs, and to substantial rework.

4.3 Case 3: Purchasing Association in the Furniture Industry

Case 3 describes coordinating quality of product master data in a purchasing association in the furniture industry, hereafter referred to as C3. The association serves as a mediator between furniture manufacturers and retailers. C3 draws up and makes framework contracts between producers and vendors, and provides procurement, marketing, financing, and IT services. C3 supports, for example, the exchange of master data between furniture suppliers and retailers. C3 employs more than 50 people and generates an annual turnover of about 50 million Euros. 700 furniture producers and dealers are members of C3.

C3's product data base contains about 100,000 master data sets. More than 1000 new master data files are imported and approx. 30,000 master data sets need to be modified per year. The organizational unit responsible for managing master data directly reports to C3's executive board. The master data team consists of 5 employees. C3's process of arranging for and agreeing on an appropriate level of master data quality is similar to the processes at C1 and C2.

- C3 sends Excel files to furniture manufacturers requesting to provide master data in a form that meets requirements of furniture retailers. About 50 percent of the suppliers fill in the templates and provide master data in Excel files. The other half provides master data sets in proprietary formats, in PDF files, or on paper.
- Similar to the situation at C1 and C2, a high percentage of all files provided by suppliers are incomplete or incorrect. This gives rise to frequent check backs. The process is labor intensive and causes high personnel cost. However, our interview partners were not able to quantify efforts needed to complete or to amend master data sets.
- C3's procurement team regularly communicates with suppliers. They attempt to clarify any inconsistencies. However, arranging for and ensuring an appropriate level of master data quality is not a top priority for procurement personnel. Master data experts are not involved in talks and negotiations with furniture manufacturers. Similarly, C3's IT department does not actively support the process of exchanging master data with suppliers.
- Unlike C1 and C2, C3 does not complement master data sets with additional attributes. C3 serves as a data broker only.
- C3 does not apply any product identification or classification standard. They neither use EDI messages for transferring master data nor data pools for enriching and amending incomplete or incorrect data sets.
- C3 does not use any automatic routines to check the quality of master data provided by suppliers before transferring master data sets to furniture retailers. Once C3's master data team considers the data sets appropriate, they transfer the data into a product data base. Product information is then made available for furniture retailers via C3's extranet, a controlled private network allowing members to gain proprietary information.

5 Findings and Discussion

In this section, we answer the research questions and discuss findings of our study. We first present the assessment of coordination tools used by the trading companies included in our case studies. We then analyze our findings using the lens of coordination theory.

5.1 Assessment of Coordination Tools used by Trading Companies

During the case study interviews, we identified a set of eight coordination tools that are applied to define and to check quality of product master data exchanged between suppliers and wholesale companies. However, as already mentioned in section 4, not each company included in our study makes use of all of these tools.

As part of the case study interviews, we asked participants to assess the tools. We invited the experts to qualitatively evaluate strengths and weaknesses of the tools. Table 1 shows the results of the assessment.

Table 1. Assessment of coordination tools for ensuring quality of product master data

Coordination Tools	*Strengths*	*Weaknesses*
Framework contract	- provides legal basis for data exchange including confidentiality agreements	- does not specify requirements for master data quality
MS Excel file	- ease of use - high adoption rate - simple editing - no IT expertise required - low set up cost	- most suppliers do not document modifications in files - complex rework required - automated validation of master data sets not feasible - files are not self-descriptive
EDI message	- standardized, tested and reliable technology - reliable data sets	- low adoption rate - EDI messages cannot easily be tailored to specific requirements - high implementation cost
Customer portal	- automatic syntax check - documentation of all operations in log files	- specific "look and feel" of each portal - insufficient ease of use - data sets are not tailored to specific needs of trading partners
Online product catalog	- all products can be mapped and made available	- trading partners must be persuaded to use product catalogs
Identification and Classification standard	- high data quality - high adoption rate - full documentation and good support	- cannot be tailored to specific requirements - low adoption rate
Data pool	- covers a wide range of products - focus on specific industries - high data quality - unified communication platform - communication processes are logged	- limited number of attributes - specific "look and feel" of each data pool - high access and usage fees - not all business partners have access
Mutual adjustment (e-mail messages, telephone calls, personal meetings)	- ease of use - flexibility	- error prone - labor-intensive - costly

5.2 Analyzing the Cases Using the Lens of Coordination Theory

We build on coordination theory to structure our findings and to reflect the insights we gained in our case studies. Malone and Crowston [27] define coordination as "the act of managing dependencies between entities and the joint effort of entities working together towards mutually defined goals". Coordination theory is the "body of principles about how activities can be coordinated, that is, about how actors can work together harmoniously" [28]. Thus, coordination theory focuses on essential questions of inter-organizational product information sharing, namely, how trading partners can work together harmoniously in order to ensure smooth and cost-efficient supply processes and how they can manage interdependencies resulting from the need to operate on high quality product master data. Another reason for selecting coordination theory as a theoretical basis is that it matches well with our research questions. Coordination theory explores which types of coordination mechanisms have which strengths and shortcomings under which conditions [27, 29, 30]. Moreover, it has been successfully applied by prior research into inter-organizational product information sharing [8, 14].

Inter-organizational coordination of master data quality is a highly complex, yet hardly systematized or automated process in the companies covered by our study. The process of check backs and providing previously missing or defective data is not standardized or automated. Procurement employees and master data staff perform these tasks on an ad hoc basis. Most subtasks are performed manually. Apparently, many suppliers and their customers make considerable efforts to achieve and to maintain an adequate level of master data quality. However, options for standardization, rationalization, and automation have not yet been fully used by the companies included in our study. These findings confirm the research results of Legner and Schemm and Otto et al. [4, 8].

Several authors [27–31] have suggested frameworks that describe basic concepts of inter-organizational coordination.

Actors denote individuals and organizational units responsible for coordination [28]. Entities responsible for the quality of product master data are centralized in all companies covered by our study. However, the process of coordinating master data quality with suppliers is decentralized. Procurement personnel and staff members of master data teams are involved. Both entities adjust incorrect data and complement incomplete data sets. This requires additional intra-organizational coordination effort. We had expected that IT departments play a crucial role in the exchange of master data sets and in ensuring and improving master data quality. However, IT departments do not actively participate in this process. IT staff run the IT infrastructure and they assist in uploading master data files into ERP systems. Yet, they are not consulted when it comes to technical solutions for improving inter-organizational coordination of master data quality.

Interdependencies are “goal-relevant relationships between the activities” [28].. According to Malone and Crowston [28] interdependencies can be analyzed in terms of common objects. Common objects in a trading relationship are trade items and product master data sets that are transferred between trading partners. However, master data are not of key importance for all actors participating in the trading process. Procurement personnel rather focus on trade items, prices, and terms and conditions. Improving quality of master data is seen as an annoying and bothersome technical task. Thus improving quality of master data does not receive the necessary attention. As mentioned before, IT personnel – who would be able to adequately assess the importance of the task – usually are not involved in inter-organizational product information sharing.

Romano [32] differentiates between several types or configurations of inter-organizational relationships. A dyadic network involves the interaction between two firms (1:1), a multiple dyadic network involves the interaction of one firm with several other firms (1:n or n:1). A many-to-many network is one where several firms interact with several other firms (m:n). The types or configurations analyzed in our study are multiple dyadic networks. This corresponds to the results of Le Dû and de Corbière [19] who found that synchronization of product information is mainly performed on a dyadic basis. Each of the three wholesale organizations has trade relations to several hundred suppliers. Quality of master data needs to be arranged, agreed on, and checked on a one-to-one level between trading partners. Obviously, the extent and complexity of this task makes it impossible to ensure an adequate quality of product information exchanged with all trading partners.

Xu and Beamon [29] define coordination mechanisms as a set of methods used to manage interdependencies between organizations. Coordination tools are specific elements of organizational action, interaction or behavior that enable inter-organizational coordination [31]. Arshinder et al. [30] suggest identifying whether a single or a combination of mechanisms are required to tackle complexities in supply chain coordination. The wholesale trading companies included in our study apply a multifaceted set of methods to coordinate master data quality with suppliers. However, Excel files are the predominant coordination tools. These files allow trading companies to accurately define quality specifications for product master data. By defining attributes and macros in spreadsheet files, wholesale trading companies attempt to standardize product master data sets requested from their suppliers. However, a considerable number of files provided by suppliers are incomplete or incorrect. This gives rise to frequent check backs. Furthermore, it is worth stressing, that our analysis of these tools revealed several weaknesses. Some of the Excel files used in the companies covered by our study are difficult to use, partly incorrect or technically outdated. This makes it more difficult for suppliers to provide master data in an appropriate quality.

Numerous incomplete and defective data sets provided by suppliers require frequent check backs. E-mail messages and telephone calls are regularly used to support mutual adjustment by data experts at trading companies and their suppliers.

In some cases, suppliers provide master data sets in EDI messages. Only a very limited set of suppliers provide master data for wholesale customers via customer portals or online product catalogs. Product identification and classification standards are used to a very limited degree in the companies covered by our study. Only one company utilizes a data pool. On the one hand, a more intensive use of standards and tools provided by electronic intermediaries, e.g., data pools and electronic catalogs, may contribute to improving product data quality in inter-organizational product information sharing and to reducing the workload for most trading partners. On the other hand, these tools obviously need to be complemented by bilateral arrangements that account for specific requirements of the particular trading partners.

Our findings confirm the results of previous research [1, 4, 5, 8, 9, 14, 18, 19, 33]. Earlier studies revealed that EDI connections, identification and classification standards, product catalogs, and data pools are not widely adopted in most industry sectors. Instead, most companies prefer Excel files and personal communication to coordinate master data quality in inter-organizational product information sharing.

These phenomena can be explained with experts' perceptions of strengths and weaknesses of the coordination mechanisms and tools described in section 5.1. Personnel in master data teams and procurement staff in the companies covered by our study either do not have sufficient knowledge of standards, EDI technology, product catalogs, or data pools or they are convinced that these resources are too cost-intensive or not flexible enough.

6 Conclusion

6.1 Implications for Practitioners

The insights we gained in our study indicate some interesting options for improvement. We believe the recommendations presented in this section might also be relevant for other organizations facing similar issues.

Only one company included in our study explicitly mentions exchange of product information in framework contracts with trading partners. This could encourage other companies to do the same to raise awareness for achieving and maintaining adequate quality of master data in inter-organizational product information sharing.

The process of coordinating quality of product information with trading partners is only scarcely systematized and automated. A higher level of standardization and automation could help increase process efficiency, improve master data quality, and reduce personnel cost. This, however, would require establishing a well-defined process that is accepted by all trading partners, specifically by master data experts. Further above we have outlined the current state of the process for coordinating product master data quality. Defining and testing an improved process that is more intensively systematized, automated, and accepted by all parties would be an essential improvement. An improved process including automated mechanisms and tools could release all trading partners from manual data entries and cost-intensive rework after having received incomplete or inappropriate master data sets.

We do not expect that such a process can be established in the short term. In the meantime, however, regular meetings of master data experts of selected suppliers and customers could help to improve the exchange of master data with at least some trading partners. Best practices that hopefully result from these forms of cooperation could then be adopted by more trading partners.

6.2 Research Opportunities

Although the nature of our study is such that no universally applicable conclusions can be drawn, our findings at least illustrate the necessity to conduct more in-depth research into coordinating master data quality in inter-organizational product information sharing. There are a number of interesting extensions for future research.

One area is to include more companies in future studies. It would be particularly interesting, to interview master data experts employed at suppliers and their contact partners in trading companies in order to gain more complete insights into benefits and downsides of applying coordination mechanisms and tools in inter-organizational master data sharing. Another option to broaden research is to explore inter-organizational coordination in other sectors (e.g., manufacturing, healthcare etc.) to identify potential differences in coordinating quality of product information.

We also encourage scholars and practitioners to design, implement and evaluate software tools that support coordinating quality of product master data in inter-organizational relations. Such tools could help to define requirements for master data to be exchanged among trading partners. Software tools could support examining the extent to which requirements are fulfilled. Some authors [17, 34–36] have proposed platforms and services for improving quality of shared data in cooperative information systems. The functionality of these tools seems to be limited and none of the tools has been developed to a product that is available on the market. However, the papers may provide an interesting starting point for the design and development of more comprehensive tools. Automated coordination tools could provide workflows for adjusting incorrect data or complementing incomplete data sets between trading partners. Once, such a tool is completed it would be highly interesting to explore whether it can help to reduce costs of coordinating quality of master data and, if so, to what extent.

References

1. Legner, C., Wende, K.: The Challenges of Inter-Organizational Business Process Design - A Research Agenda. In: ECIS 2007 Proceedings (2007)
2. de Corbière, F.: Interorganizational Information Systems and Data Quality Improvement: The Case of Product Information in the French Large Retail Industry. In: ICIQ 2007 Proceedings (2007)
3. Loshin, D.: Master Data Management. Morgan Kaufmann, Amsterdam (2008)
4. Otto, B., Abraham, R., Schlosser, S.: Toward a Taxonomy of the Data Resource in the Networked Industry. In: ISSL 2014 Proceedings (2014)
5. Dalmolen, S., Moonen, H., van Hillegersberg, J.: Industry-wide Inter-organizational Systems and Data Quality: Exploratory findings of the use of GS1 standards in the Dutch retail market. In: AMCIS 2015 Proceedings (2015)
6. Wang, R.Y., Strong, D.M.: Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems 12, 5–33 (1996)
7. Otto, B., Österle, H.: Corporate Data Quality. Springer, Heidelberg (2016)
8. Legner, C., Schemm, J.: Toward the Inter-organizational Product Information Supply Chain - Evidence from the Retail and Consumer Goods Industries. Journal of the Association for Information Systems 9, 119–150 (2008)
9. Madlberger, M.: Can data quality help overcome the penguin effect? The case of item master data pools. In: ECIS 2011 Proceedings (2011)
10. Haug, A., Stentoft Arlbjørn, J.: Barriers to master data quality. Journal of Enterprise Information Management 24, 288–303 (2011)
11. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26, xiii–xxiii (2002)
12. vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In: ECIS 2009 Proceedings (2009)
13. Nakatani, K., Chuang, T.-T., Zhou, D.: Data Synchronization Technology. Communications of the Association for Information Systems 17, 962-996 (2006)
14. de Corbière, F.: Data Quality and Interorganizational Information Systems: The Role of Electronic Catalogues. In: AMCIS 2009 Proceedings (2009)

15. Hüner, K.M., Schierning, A., Otto, B., Österle, H.: Product data quality in supply chains. The case of Beiersdorf. Electronic Markets 21, 141–154 (2011)
16. Falge, C., Otto, B., Österle, H.: Data Quality Requirements of Collaborative Business Processes. In: HICSS 2012 Proceedings (2012)
17. Cai, Y., Shankaranarayanan, G.: A Data Quality Assurance Model in the B2B Networked Environment. In: AMCIS 2004 Proceedings (2004)
18. de Corbière, F., Rowe, F.: Understanding the Diversity of Interconnections between IS: Towards a New Typology of IOS. In: ECIS 2010 Proceedings (2010)
19. Le Dû, A.-C., de Corbière, F.: IQ as an enabler of the green and collaborative supply chain. In: ICIQ 2011 Proceedings (2011)
20. Morgan, D.L.: Successful Focus Groups: Advancing the State of the Art. SAGE, Thousand Oaks (1993)
21. Rosemann, M., Vessey, I.: Toward Improving the Relevance of Information Systems Research to Practice: The Role of Applicability Checks. MIS Quarterly 32, 1–22 (2008)
22. Chiarini Tremblay, M., Hevner, A.R., Berndt, D.J.: The Use of Focus Groups in Design Science Research. In: Hevner, A., Chatterjee, S. (eds.) Design Research in Information Systems. Theory and Practice, pp. 121–143. Springer, Boston (2010)
23. Eisenhardt, K.M., Graebner, M.E.: Theory Building from Cases: Opportunities and Challenges. Academy of Management Journal 50, 25–32 (2007)
24. Yin, R.K.: Case study research. Design and methods. SAGE, Los Angeles (2014)
25. Mayring, P., Fenzl, T.: Qualitative Inhaltsanalyse. In: Baur, N., Blasius, J. (eds.) Handbuch Methoden der empirischen Sozialforschung, pp. 543–556. Springer, Wiesbaden (2014)
26. TecAlliance GmbH: TecDoc catalog, https://www.tecalliance.net/en/products/tecdoc-catalogue-portal/ (Accessed: 20.06.2106)
27. Malone, T.W., Crowston, K.: The Interdisciplinary Study of Coordination. ACM Computing Surveys (CSUR) 26, 87–119 (1994)
28. Malone, T.W., Crowston, K.: What is coordination theory and how can it help design cooperative work systems? In: CSCW 1990 Proceedings (1990)
29. Xu, L., Beamon, B.M.: Supply Chain Coordination and Cooperation Mechanisms: An Attribute-Based Approach. Journal of Supply Chain Management 42, 4–12 (2006)
30. Arshinder, K., Kanda, A., Deshmukh, S.G.: A Review on Supply Chain Coordination: Coordination Mechanisms, Managing Uncertainty and Research Directions. In: Choi, T.-M., Cheng, T.E. (eds.) Supply Chain Coordination under Uncertainty, pp. 39–82. Springer, Heidelberg (2011)
31. Alexander, E.R.: Interorganizational Coordination: Theory and Practice. Journal of Planning Literature 7, 328–343 (1993)
32. Romano, P.: Co-ordination and integration mechanisms to manage logistics processes across supply networks. Journal of Purchasing and Supply Management 9, 119–134 (2003)
33. Ofner, M.H., Otto, B., Österle, H.: Integrating a data quality perspective into business process management. Business Process Management Journal 18, 1036–1067 (2012)
34. Mecella, M., Scannapieco, M., Virgillito, A., Baldoni, R., Catarci, T., Batini, C.: Managing Data Quality in Cooperative Information Systems. In: CoopIS 2002 Proceedings (2002)
35. Scannapieco, M., Virgillito, A., Marchetti, C., Mecella, M., Baldoni, R.: The DaQuinCIS architecture. Information Systems 29, 551–582 (2004)
36. Becker, J., Matzner, M., Müller, O., Winkelmann, A.: Towards a Semantic Data Quality Management - Using Ontologies to Assess Master Data Quality in Retailing. In: AMCIS 2008 Proceedings (2008)

Integrated Optimization of IT Service Performance and Availability Using Performability Prediction Models

Sascha Bosse, Hendrik Müller, and Klaus Turowski

Very Large Business Applications Lab
Faculty of Computer Science, Otto-von-Guericke University Magdeburg
{sascha.bosse|hendrik.mueller|klaus.turowski}@ovgu.de

Abstract. Optimizing the performance and availability of an IT service in the design stage are typically considered as independent tasks. However, since both aspects are related to one another, these activities could be combined by applying performability models, in which both the performance and the availability of a service can be more accurately predicted. In this paper, a design optimization problem for IT services is defined and applied in two scenarios, one of which considers a mechanism in which redundant components can be used both for failover as well as handling overload situations. Results show that including such aspects affecting both availability and performance in prediction models can lead to more cost-effective service designs. Thus, performability prediction models are one opportunity to combine performance and availability management for IT services.

Keywords: IT Service Management, Availability Management, Capacity Management, Performability Modeling, Redundancy Allocation Problem

1 Introduction

IT service providers are faced with the challenge of designing high-quality and cost-effective IT systems in order to stay competitive [1]. In particular, degraded quality of service may cause the violation of service level agreements, leading to penalty costs and loss of reputation for the service provider [2]. Two of the most crucial quality aspects of an IT service are performance and availability [3]. In order to ensure that a service achieves the desired quality level at a minimum of costs, capacity management for performance as well as availability management have to be performed carefully [3].

However, managing the quality and costs of an IT service is difficult as these are complex systems with specific requirements, an architectural mix of diverse hardware and software components as well as skewed workload [1], [4]. In order to have accurate estimates about service quality, an IT service is usually tested before its deployment [5], [6]. Nevertheless, there are decisions to be made in the service design stage and their correction in the deployment or operation stage may be very costly [7], [8]. This applies especially for architectural decisions which have the greatest impact on quality attributes [9].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Bosse, S.; Müller, H.; Turowski, K. (2017): Integrated Optimization of IT Service Performance and Availability Using Performability Prediction Models, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 76-90

Therefore, performance and availability modeling techniques are effective in the service design stage, as they provide a first estimate about the service's quality level [1], [5], [10]. On the basis of these models, different design alternatives can be compared to minimize service costs [1], especially if alternatives are recommended automatically [4]. For that purpose, optimization approaches can be applied. For instance, in [3] a heuristic for service-component allocations is presented to support capacity management. Other examples include optimizing the redundancy design for availability management by defining redundancy allocation problems (RAP), e.g. in [11], [12].

Despite its advantages, a recent report of Gartner comes to the conclusion that most enterprises under-invest in performance modeling for capacity management [13]. One reason for that may be that most existing approaches for capacity management consider performance in isolation, disregarding dependencies to other quality attributes such as reliability, availability, or scalability [14], which can be subsumed under the term dependability [15]. Thus, performance prediction models overestimate the ability of a service to satisfy its consumers [16]. On the other hand, applying pure dependability approaches will often lead to conservative quality estimates so that cost saving potentials are not addressed [16]. Hence, combined models for availability and performance prediction could depict an opportunity to integrate capacity and availability management in order to achieve a cost-effective and high-quality service design.

For approaches in which both performance and dependability aspects are considered, the term performability modeling can be used. While performability prediction models have been developed to analyze a single system, their possible impact on design optimization considering a large variety of systems has not been researched. Therefore, the question remains if an optimization approach on the basis of performability models would allow for higher cost-effectiveness of IT services than if isolated approaches are applied. In order to answer this question, an optimization problem is formulated in this paper for the capacity and redundancy design of an IT service. A simple Petri net performability model is used to estimate the quality and costs of a service design. In an illustrative example, it can be demonstrated that performability models provide the flexibility to include mechanisms affecting both performance and availability of a service such as an elastic standby redundancy derived from cloud computing concepts, which is not only used in failure but also in overload situations. This leads to a more effective peak-load handing [17] and, thus, to design suggestions with less costs while satisfying performance and availability service level objectives.

2 Related Work

In this section, relevant basics of capacity and availability management as well as performability modeling are introduced as the basis for the optimization problem.

2.1 Capacity Management and Performance Models

The capacity management process is an important part of the design stage in IT service management frameworks such as the ITIL [18]. Its purpose is to ensure that an

IT service meets the current and future performance-related requirements in a cost-effective manner [18]. Related terms are capacity planning and software performance engineering. The process of capacity management relies on performance models estimating the relation between input (workload model, design alternatives) and output variables (e.g. response time, throughput, or utilization) [7]. Performance models can be classified into measurement- and model-based approaches [5].

Measurement-based performance models require existing systems to be observed for at least some of the possible design alternatives [19]. The properties of non-existing systems can be estimated by using machine learning methods such as support vector machines or random forests [20] (black-box approach). Nonetheless, these performance models often lack transferability to other problem classes [5]. In model-based approaches, analytical performance models are used to describe the relationship between input and output variables (white-box approach). Examples are queuing networks [7], queuing Petri nets [4], (generalized) stochastic Petri nets [10], or stochastic reward nets [16]. However, constructing these models requires insight into the performance characteristics of the used system components, which may not be accessible especially for third-party software components [21]. Besides the pure measurement- or model-based approaches, a grey-box approach can be applied, e.g. by using gained knowledge of existing components and design patterns [5] to combine low-level measurements and high-level analytical models [8].

In order to evaluate performance models, the following techniques can be used: prototyping, testing, simulation, or analytical evaluation [5]. While prototyping and testing provide more accurate results, these approaches are also very costly and can only be applied in later lifecycle phases [6]. On the other hand, analytical and simulation techniques can be applied in the design phase of the service lifecycle [5]. Analytical evaluation is faster, but some aspects of the system performance may not be considered (e.g. G/G/n queues). In this case, simulation approaches are more effective [4], [5], [7] which also allow for dynamic analysis of performance metrics [22].

After performance models can be constructed and evaluated, many possible design alternatives should be analyzed in order to find a suitable cost-performance tradeoff [1]. Since these alternatives should be compared quickly [7], an optimization approach can be applied. For instance, in [3] an optimization problem is defined to find the minimum number of computing nodes for a set of IT services, and a component-allocation heuristic is developed to identify (sub)optimal solutions.

2.2 Availability Management and the Redundancy Allocation Problem

Similar to capacity management, availability management is an IT service design process aiming at meeting the agreed availability requirements at minimum costs [18]. A related term is software reliability engineering, which is, however, more focused on single software components than on complex systems of (third-party) components [23]. As in capacity management, black- and white-box availability models can be constructed and evaluated by different methods. White-box approaches for availability modeling include combinatorial, state-space-based, or hierarchical techniques [24].

Basically, the availability of a system can be increased by applying fault removal, fault prediction, fault prevention, and fault tolerance approaches [15] while the latter two can be considered in the design phase [25]. However, preventing faults by increasing component reliabilities is limited [26] and, thus, fault-tolerance has to be applied in order to achieve high-availability, e.g. by introducing redundancy mechanisms [27]. Redundancy means to provide additional components with equal functions to cover faults of the original components. Redundant components may be purchased from another manufacturer or developed by different teams to decrease the probability of common-cause failures (heterogeneous redundancy). Additionally, active and passive (or standby) redundancy can be distinguished.

A component in active redundancy is ready to instantly takeover for a defect component. Passive components are in a lower state of readiness and need some time for takeover. Depending on the state and the time to activation, standby components can be classified into cold-, warm-, and hot-standby [28]. With increasing time to activation, usually failure rate and operational costs of a passive component are decreased [27]. In order to decide which components may require redundancy mechanisms and which type of redundancy has to be used, a redundancy allocation problem (RAP) can be defined. A RAP is a NP-hard optimization problem [29] and is often characterized by the following aspects:

1. A system consists of n required subsystems.
2. In each subsystem, a number of components can be operated in redundancy.
3. Components fail and recover according to random time distributions.
4. For each subsystem, a number of components has to be identified so that availability is maximized or costs are minimized subject to constraints.

In recent years, RAP definitions with increasing complexity have been developed to include characteristics of IT systems. For instance, standby redundancy and its effects have been considered in [12], [28]. RAP are usually solved using (meta-)heuristics [30] such as genetic algorithms, e.g. in [11], [12], [28].

2.3 Performability Modeling and Design Optimization

While isolated performance and availability models are still subject of research, several approaches try to combine both dependability and performance aspects as there are significant relations between those. One reason for this is that availability is not defined as the absence of failures, but as the probability of success [31]. Thus, unavailability is not only caused by failures but also by overload situation which lead to request rejections [19]. Furthermore, when response times increase, users will abort waiting for a reply and will consider the service as unavailable [7]. On the other hand, component faults may lead to degraded capacity of a service.

In order to consider these aspects in a more accurate model of performance and availability, the concept of performability models can be applied [31]. Examples can be found in [5], [10], [16], [19]. However, there is a lack of design optimization approaches that are based on performability models although some RAP definitions are able to address performance aspects superficially since system capacity is formed by the components operated in active redundancy. For instance, multiple component states have been introduced to model performance degradation of components. In order to compute the system performance from component states, the loss of load probability (LOLP) index is used as a measure for unavailability (e.g. in [32]) which has been originally defined for electrical power systems as the probability that demand exceeds capacity. In an IT system, throughput is a related concept [33]. In a parallel subsystem of (active) redundant components, the throughput can be computed as the sum of the components' throughputs. The throughput of a series system is defined as the minimum of the subsystems' throughputs. In order to estimate the availability of the system, the load of the system is set in relation to its throughput.

The disadvantage of this approach is that performance is not considered in detail so that, for instance, response times cannot be analyzed. Furthermore, all subsystems have the same load at every time. In an IT service, however, a subsystem may be visited by a request more than once and the processing time in a subsystem can differ significantly [7]. In addition to that, the LOLP approach leads only to estimates about the mean utilization of components which makes it difficult to get accurate estimates for operational costs which may depend strongly on current utilization as e.g. the power consumption of a CPU. Thus, this approach is not suitable for an effective design optimization for a combined capacity and availability management.

3 The Redundancy and Capacity Allocation Problem (RCAP)

Our study of related work revealed that no suitable design optimization problem has been defined on the basis of performability models. In order to investigate the feasibility of performability modeling for design optimization, a new optimization problem is presented in this section on the basis of the RAP. The objective of the optimization problem is to minimize the costs of an IT service while satisfying mean response time and availability constraints. Although the capacity of a service is determined by the components operated in active redundancy, the problem is named redundancy and capacity allocation problem (RCAP) to underline the differences to classical RAP approaches in which performance aspects are not sufficiently addressed.

In order to demonstrate the opportunities of performability models, two different strategies for components in passive redundancy are considered: in the classical standby case, passive components are activated if and only if another component has failed. In the elastic standby scenario, these components are activated only in case of overloads, which may be caused by load peaks as well as component failures.

3.1 Assumptions

An IT service is consumed by customers that send independent, comparable requests. In order to respond to a request, a sequence of operations has to be performed. An operation is executed in a subsystem consisting of functional equivalent components. Each of these components can serve a number of operations concurrently and the processing time depends on the capacity of the component.

A component can have three states: active, standby (passive), and failed. Initially, a component is either active or standby. In active mode, a component fails and is recovered again after random time intervals. A component does not provide capacity if it is failed and cannot fail if it is in standby mode. If a standby component has to be activated, a random time elapses before the component is set to active mode and may fail. In addition to component failures, the whole service can also be affected by a failure, which leads to the rejection of all requests currently processed (disaster case). All failure, recovery and activation times as well as inter-arrival and processing times can be described by parametric random distributions.

If all servers are busy in a subsystem, operations to be processed are queued until a server becomes available (FIFO strategy). An admission control can be enabled to avoid unnecessary processing if a user will likely abort waiting for a reply by automatically rejecting requests after a certain time in the system (timeout).

The costs of the service are the sum of capital and operational costs. While the former is determined by the acquisition costs for components, the latter costs are compound of the costs arising for recovery activities and the components' operational costs that depend linearly on their utilization levels. Another linear dependency is assumed between components and system operational costs, which is characterized by the power usage effectiveness (PUE) that ranges in practice from 1.2 to 3.0 [34].

3.2 Problem Definition

A redundancy and capacity allocation problem (RCAP) is characterized by a timestep Δt, a factor PUE as well as by random variables R, TTF and TTR describing inter-arrival times, times to failure and times to recover for an IT service. Furthermore, a number of operations o and subsystems n have to be defined as well as a function ops mapping operations to subsystems. Another function S maps the operations to random variables describing the standard service time. In addition to that, a timeout to can be defined for admission control.

A set of components is associated with each subsystem. Each component i is characterized by its number of servers for concurrent operations s_i, a service time norm factor f_i, random distributions TTF_i and TTR_i, a random distribution TTA_i for its time to activation, its acquisition costs ac_i, recovery costs rec_i as well as operational costs for Δt in standby state $standby_i$, in idle mode $idle_i$ and if fully utilized $full_i$.

A solution candidate x to the RCAP consists of $2n$ sets of component selections describing the components to be operated in active as well as passive redundancy for each subsystem. Thus, the optimization problem can be defined as follows:

$$\min_x C(x,t) \text{ s.t. } A(x) \geq A_0 \wedge RT(x) \leq RT_0 \tag{1}$$

Since the costs of a service C depends on the time interval to be considered, a parameter t has to be defined as well as the service level objectives for availability A_0 and response time RT_0.

An overview of all required parameters to define a RCAP are presented in Table 1.

Table 1. Parameter overview for the RCAP

Problem Parameters	Description
$\Delta t \in \mathbb{R}$	Timestep for operational costs definition
$PUE \in \mathbb{R}$	Power usage effectiveness of the IT service
$R: \mathbb{R} \to \mathbb{R}$	Distribution of request inter-arrival times
$TTF, TTR: \mathbb{R} \to \mathbb{R}$	Distribution of time to failure/recovery for the service
$o, n \in \mathbb{N}$	Number of required operations and subsystems
$ops: \mathbb{N} \to \mathbb{N}$	Function mapping operations to subsystems
$S: \mathbb{N} \to (\mathbb{R} \to \mathbb{R})$	Function mapping operations to standard service time distributions
$to \in \mathbb{R}$	Timeout after which admission control is applied
$s_i \in \mathbb{N}$	Number of servers in a component *i*
$f_i \in \mathbb{R}$	Service time norm factor for a component *i*
$TTF_i, TTR_i, TTA_i: \mathbb{R} \to \mathbb{R}$	Distributions of time to failure, recovery, and activation for a component *i*
$ac_i, rec_i \in \mathbb{R}$	Acquisition and recovery costs of a component *i*
$standby_i, idle_i, full_i \in \mathbb{R}$	Operational costs per timestep in standby and idle mode as well in full operation for a component *i*

3.3 A Simple Performability Prediction Model

On the basis of the RCAP definition, a simple performability prediction model is developed that can be used to estimate availability, mean response time, and costs of the solution candidates. Due to their modeling power, a solution candidate is modeled in a generalized stochastic Petri net (GSPN), cf. e.g. [35]. In order to consider arbitrary random variables, these models are evaluated by Monte Carlo simulation.

Thus, a GSPN has to be automatically generated for each solution candidate. First, an arrival transition is defined that generates request tokens. To each subsystem, a queue place is assigned with infinite capacity for collecting incoming requests. Another place collects all rejected requests for the service. For a subsystem, a place and a corresponding processing transition are created for each server of the assigned components and are connected to the queue of the subsystem performing the next operation. The subsystem queue is connected to the server places by immediate transitions which are only active if the component is working. In the case that the corresponding component fails, another immediate transition fires tokens to the rejection place instead. To all transitions, random firing time distributions are assigned according to the component's definition.

In
Figure **1**, an example GSPN of a single subsystem of one active component with two servers is presented. Activation functions (dashed rectangles) are used to prevent requests to be rejected if the component is working or to be processed if it is failed. Request are created by the arrival transition and collected in the queue place. Depending on the component state (failed/working cycle below), requests are rejected or processed and assigned to the next operation queue.

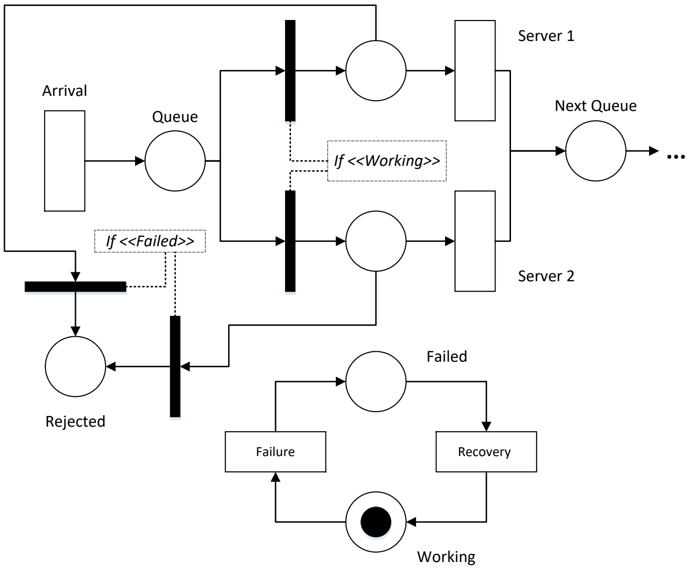


Figure 1. GSPN processing model example with an active component

In order to implement the operation sequence as well as the admission control, tokens are distinguishable (colored Petri net) by their arrival time and the current operation. Based on this information, all tokens that have been longer in the system than the defined timeout are moved to the rejection place. Additionally, a subsystem model such as presented above can map more than one operation since firing times depend on the current operation of the token. After all operations are performed, the token is moved to a place in which completed requests are stored.

While the state of an active component can be modeled by a simple failed/working cycle, for standby components, a standby place is introduced to indicate the current state of the passive component as displayed in Figure 2.

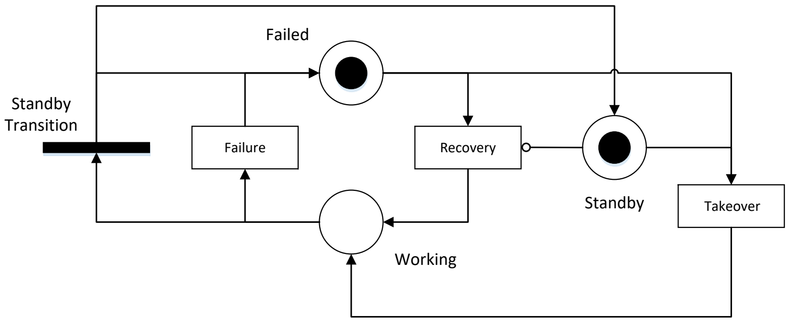


Figure 2. GSPN state model example of a passive component

In standby mode, the state of the component is equivalent to a failure, but recovery is disabled. To the takeover and standby transitions, activation functions are assigned depending on the subsystem characteristics and the chosen standby strategy. In case of the classical standby, the takeover transition is activated if an active component fails. If the active components are recovered, the standby transition is activated and the passive component changes to standby mode. In elastic standby, takeover is activated if the subsystem queue is not empty and all working servers are busy (overload situation). The component is deactivated again if none of its servers are busy.

For each completed request, the response time is the difference between its completion and its arrival time. The mean response time of a solution candidate is computed from all request response times at the end of a simulation run. Availability is defined as the ratio of completed requests to arrived requests minus those in completion. The costs are the sum of acquisition, recovery, and operational costs. Impulse rewards for recovery costs are assigned to recovery transitions. The utilization of a component is computed after each timestep Δt and is used to compute its current operational costs c_i linearly based on values $idle$ and $full$. This value is used in order to estimate the service operational costs C_{op} up to a time *t*:

$$c_i(t) = \begin{cases} standby_i & \text{, if} \ll Failed \gg \\ idle_i + u_i(t) \cdot full_i & \text{, otherwise} \end{cases}$$
$$C_{op}(x,t) = \frac{PUE}{\Delta t} \cdot \int_0^t \left(\sum_{i \in x} c_i(\tau) \right) \mathrm{d}\tau \tag{2}$$

By simulating the behavior of the solution candidate in several independent runs, mean values for response time, availability, and costs can be computed to characterize a solution candidate with statistical significance.

3.4 A Genetic Algorithm to Solve the RCAP

The RCAP defines an infinite search space that has to be efficiently explored for identifying (sub)optimal solutions. Therefore, a genetic algorithm is defined, which is characterized by problem-specific operations for evaluating a solution's fitness, encoding a solution, initializing random solutions, altering (mutation) and recombining solutions as well as selection and termination criteria.

The suitability of a solution to solve the defined RCAP is characterized by its fitness to be maximized. According to [11], a penalty function is used to avoid infeasible solutions in the end. Since costs has to be minimized, the fitness is defined as the negative costs minus penalty:

$$f(x) = -C(x) - \delta \left(\frac{\max\big(0, A_0 - A(x)\big)}{NFT_A} + \frac{\max(0, RT(x) - RT_0)}{NFT_{RT}} \right) \tag{3}$$

In order to scale the penalty, pre-defined near-feasibility thresholds (NFT) are used as well as a factor δ, which is the difference between the minimal cost of all solutions and the minimal cost of feasible solutions (cf. [36]).

For the encoding, a solution candidate is represented by two vectors for the active and the passive configuration. Each vector consists of n integer vectors corresponding to the indices of the defined components for a subsystem. An example encoding for three subsystems may look as follows:

$$\Big(\big((2,1),(1),(2,2)\big),\big((\ ),(1),(\ )\big)\Big)$$

In this solution candidate, two different components are operated in active redundancy in the first subsystem (of the second and the first type in this subsystem), but no component is operated in passive redundancy (empty bracket in the second vector). In the second subsystem, however, two components of the first type in this subsystem are used whereas one is in active and the other in passive redundancy.

Solution candidates are generated by choosing random components for the subsystems. The number of components in a subsystem is randomly chosen according to a pre-defined distribution. If a solution is to be mutated, a subsystem is randomly chosen and a random operation is applied to active or passive sets: with 20% probability, a component choice is exchanged. A component is either added or removed with 40% probability. The remove operation is disabled if a subsystem consists only of one active or zero passive components. For the recombination of two solutions, separate uniform crossovers are applied for the active as well as the passive configuration. Thus, subsystem configurations are exchanged between both solution candidates.

In the genetic algorithm, first a number μ solutions are generated and their fitness is evaluated. In the main loop, λ solutions ($\lambda > \mu$) are created by recombination (whereas fitter solutions are more frequently recombined) and mutation is applied with a certain probability p_{mut}. After the fitness of the new solution candidates is evaluated, μ solutions are selected from those by performing tournament selection for the next loop (generation). In tournament selection, μ tournaments of t solutions are performed in which the kth fittest solution is selected with probability $p_t(1-p_t)^{k-1}$. The algorithm results in the fittest solution after $maxGen$ generations.

4 An Illustrative Example: Sales and Distribution

The advantages of the performability optimization approach are demonstrated in a small and illustrative example describing an IT service with eight operations. For that purpose, the GA, model generation and simulation have been implemented in the java-based simulation framework *AnyLogic 6.8.2*. The simulation results have been verified by comparing test instances with results from Markov-based queuing and availability models.

The operations to be performed describe a standard business process for a delivery from stock scenario with prior lead generation and subsequent material requirements check. Therefore, five different subsystems are involved in this process (cf. Table 2). In order to define service times, data from a standard benchmark for the sales and distribution modules of SAP ERP business applications has been used. In this context, the capacity of components can be measured in the normalized metric SAPS (SAP Application Performance Standard), where 100 SAPS equal 2,000 fully processed order line items per hour. Standard service times have been computed from mean response times of various benchmark results that had been performed over the last ten years and are presented in Table 2. Using the reciprocal of these values as a rate, exponential random variables have been defined for the standard service time of each operation. These times are valid for the mean processing power of all analyzed components which is 5332.62 SAPS.

Table 2. Operations, subsystems, and standard service times of the illustrative example

Operation	Subsystem	Mean standard service time in ms
Maintain Leads	Customer Relationship Management (CRM)	2,451.65
Create Sales Order	Sales and Distribution (SD)	1,415.69
Create Outbound Delivery	Logistics Execution (LE)	1,309.62
Display Sales Order	Sales and Distribution (SD)	883.65
Change Outbound Delivery	Logistics Execution (LE)	1,101.17
Create List of Sales Orders	Sales and Distribution (SD)	1,923.90
Create Billing Document	Billing (B)	2,121.83
Display Stock/Requirements Situation	Production Planning and Control (PP)	557.71

In each subsystem, three possible components can be chosen that are presented in Table 3. The norm factor for the service times is computed by relating the SAPS per server to the mean value of the standard service times. For the computation of operational costs, an energy price of 0.2814 €/kWh is assumed.

Table 3. Available components for the subsystems

Label	*ac* in €	Capacity in SAPS	*s*	*f*	*standby* in ct/h	*idle* in ct/h	*full* in ct/h
A	2,494	11,500	8	3.70	0.02	3.83	6.22
B	21,580	150,000	72	2.57	0.28	2.43	15.67
C	4,669	43,000	16	1.99	0.14	1.58	5.52

All components share the same transition times: time to failure is exponentially distributed with a mean value of 8,760h, time to recover normally distributed with mean 1.73h and a variance of 0.5h and time to activation normally distributed with mean 30s and variance 5s (cold standby). The simulation is run for three years to map depreciation of acquisition costs (26,280h). A *PUE* of 2.0 is assumed, meaning that for every unit of energy consumed by the components a unit of energy is consumed for supporting systems. For an arrival rate of 18,000 requests per hour (exponential distribution of inter-arrival times), the mean response time should be lower than 30s and the availability should be at least 99.95%. The timeout for requests is 100s.

For both the classical and the elastic standby scenario, ten iterations of the genetic algorithm are performed. In each of these iterations, ten generations are conducted with $\mu = 30, \lambda = 60, p_{mut} = 0.2, t = 4, p_t = 0.9, NFT_A = 0.0001$ and $NFT_{RT} = 5s$. Solution candidates are initialized by choosing a normally distributed number of components in each subsystem (mean 3, variance 2 for active and mean 1, variance 0.5 for passive components). In order to evaluate a solution, 25 simulation replications are performed. In Table 4, the mean and standard deviation of the fitness values in the ten iterations are presented. It can be stated that in the elastic standby case solutions with less cost are identified on average.

Table 4. Fitness statistics for ten iterations in both scenarios

	Classical	Elastic
Mean fitness	-215,180.52	-128,415.66
Standard deviation of fitness	54,668.21	49,718.86

A comparison of the fittest solutions of both scenarios over all ten iterations is given in Table 5, using the labels of Table 3 to indicate recommended components for the subsystems presented in Table 2. For each subsystem, the recommended components in active (a) and passive (p) mode are displayed. While availability and response time are comparable in the elastic case, the costs of the service are significantly reduced in comparison to the classical standby scenario. The main reason for this is that fewer components are used in active mode.

Table 5. Fittest individuals in the classical (above) and the elastic standby scenario (below)

CRM		**SD**		**LE**		**B**		**PP**		***A***	***RT***	***C***	***-f(x)***
a	p	a	p	a	p	a	p	a	p	0.9…	in s	in €	
CCAA	A	CCCC		ACC		CC	C	AAAC		99407	25.18	119,041	131,370
AC	C	CC	C	CC		C	C	C	C	99490	24.60	84,231	84,233

These results demonstrate that a classical optimization approach will result in higher costs since capacity for overload situations has to be provided by the active components and standby components are only used in failure cases. In the elastic scenario, however, some of the needed capacity for overload situations can be provided by the standby components while failover is still applicable. Thus, considering a mechanism affecting both availability and performance in an integrated model leads to a more cost-effective service design.

5 Conclusion

In design optimization for IT services, most approaches rely on isolated models for performance or availability estimation. However, defining optimization problems on the basis of performability models considers design mechanics affecting both capacity and availability of an IT service. In this paper, it could be demonstrated that addressing these aspects in an integrated prediction model can lead to more cost-effective design suggestions than if performance and availability optimization are considered separately. Hence, this paper reveals some of the potential of integrated availability and capacity management for IT service management.

Although the developed simulation approach provides accurate results, the runtime of the optimization iterations may exceed several hours due to model stiffness, i.e. the great difference between transition rates of performance and availability state changes in a monolithic model [16]. This clearly affects the scalability of the approach to solve larger problem instances. One solution to this problem without affecting accuracy could be the massive parallelization of the independent simulation replications as well as of fitness evaluations. In this context, the scalability of the approach should be tested in large-scale problem instances in which service and transition times are not exponentially distributed to demonstrate the suitability of the simulation approach.

On the other hand, some assumptions made to demonstrate the feasibility of the applied approach have to be overcome in future work to increase applicability. This includes the assumption of the linear dependency between utilization and operational costs [37] as well as the use of mean response time as an objective while service level agreements normally include percentile guarantees. However, the developed approach can easily be adapted to result in arbitrary percentiles for response time. Furthermore, mean values of replications are used for fitness evaluation of solution candidates. Instead, confidence interval boundaries can be reported which would introduce uncertainty in the optimization. Another limitation is the fixed capacity of the defined components which may not reflect elastic and pay-per-use cloud computing services. This fact may be addressed by defining component capacity as a dynamic range with according costs. Additionally, the defined prediction model is not covering all aspects influencing response time and availability, for instance, operator errors which are reported to be one of the major causes for unavailability of IT systems [38]. Considering these limitations in future work as well as introducing additional quality aspects such as security could lead to even more realistic quality and cost estimates and, thus, to better service designs.

References

1. Almeida, V.A., Menascé, D.A.: Capacity planning an essential tool for managing Web services. IT professional. 4, 33–38 (2002).
2. Emeakaroha, V.C., Netto, M.A.S., Calheiros, R.N., Brandic, I., Buyya, R., De Rose, C.A.F.: Towards autonomic detection of SLA violations in Cloud infrastructures. Future Generation Computer Systems. 28, 1017–1029 (2012).
3. Roy, N., Dubey, A., Gokhale, A., Dowdy, L.: A capacity planning process for performance assurance of component-based distributed systems. In: ACM SIGSOFT Software Engineering Notes. pp. 259–270. ACM (2011).
4. Kounev, S.: Performance modeling and evaluation of distributed component-based systems using queueing petri nets. IEEE Transactions on Software Engineering. 32, 486–502 (2006).
5. Becker, S., Koziolek, H., Reussner, R.: The Palladio component model for model-driven performance prediction. Journal of Systems and Software. 82, 3–22 (2009).

6. Nambiar, M., Kattepur, A., Bhaskaran, G., Singhal, R., Duttagupta, S.: Model Driven Software Performance Engineering: Current Challenges and Way Ahead. ACM SIGMETRICS Performance Evaluation Review. 43, 53–62 (2016).
7. Menasce, D.A., Almeida, V.A., Dowdy, L.W., Dowdy, L.: Performance by design: computer capacity planning by example. Prentice Hall Professional (2004).
8. Terlit, D., Krcmar, H.: Generic Performance Prediction for ERP and SOA Applications. In: Proceedings of the 18th European Conference on Information Systems (ECIS) (2011).
9. Williams, L.G., Smith, C.U.: Performance evaluation of software architectures. In: Proceedings of the 1st international workshop on Software and performance. pp. 164–177. ACM (1998).
10. Bosse, S., Schulz, C., Turowski, K.: Predicting Availability and Response Times of IT Services. In: Proceedings of the 22nd European Conference on Information Systems (ECIS). , Tel Aviv, Israel (2014).
11. Coit, D.W., Smith, A.E.: Reliability Optimization of Series-Parallel Systems Using a Genetic Algorithm. IEEE Transactions on Reliability. 45, 254–266 (1996).
12. Tokuno, K., Yamada, S.: Markovian availability modeling for software-intensive systems. International Journal of Quality & Reliability Management. 17, 200–212 (2000).
13. Head, I., Govekar, M.: Market Guide for Capacity Management Tools. Gartner (2015).
14. Williams, L.G., Smith, C.U.: PASA SM: a method for the performance assessment of software architectures. In: Proceedings of the 3rd international workshop on Software and performance. pp. 179–189. ACM (2002).
15. Laprie, J.-C.: Dependable Computing: Concepts, Limits, Challenges. In: 25th IEEE International Symposium on Fault-Tolerant Computing. pp. 42–54. , Pasadena, CA, USA (1995).
16. Ma, Y., Han, J.J., Trivedi, K.S.: Composite performance and availability analysis of wireless communication networks. IEEE Transactions on Vehicular Technology. 50, 1216–1223 (2001).
17. Ranjan, R., Zhao, L., Wu, X., Liu, A., Quiroz, A., Parashar, M.: Peer-to-peer cloud provisioning: Service discovery and load-balancing. In: Cloud Computing. pp. 195–217. Springer (2010).
18. Hunnebeck, L.: ITIL Service Design 2011 Edition. The Stationery Office, Norwich, UK (2011).
19. Woodside, M., Franks, G., Petriu, D.C.: The future of software performance engineering. In: Future of Software Engineering, 2007. FOSE'07. pp. 171–187. IEEE (2007).
20. Venkataraman, S., Yang, Z., Franklin, M., Recht, B., Stoica, I.: Ernest: efficient performance prediction for large-scale advanced analytics. In: 13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16). pp. 363–378 (2016).
21. Brebner, P.C.: Performance modeling for service oriented architectures. In: Companion of the 30th international conference on Software engineering. pp. 953–954. Springer (2008).

22. Jewell, D.: Performance Modeling and Engineering. Presented at the (2008).
23. Vouk, M.A.: Software reliability engineering. In: Annual Reliability and Maintainability Symposium (2000).
24. Trivedi, K., Ciardo, G., Dasarathy, B., Grottke, M., Matias, R., Rindos, A., Vashaw, B.: Achieving and Assuring High Availability. In: Nanya, T., Maruyama, F., Pataricza, A., and Malek, M. (eds.) 5th International Service Availability Symposium (ISAS). pp. 20–25. Springer Verlag Berlin Heidelberg, Tokyo, Japan (2008).
25. Garg, H., Rani, M., Sharma, S.P., Vishwakarma, Y.: Bi-objective optimization of the reliability-redundancy allocation problem for series-parallel system. Journal of Manufacturing Systems. 33, 335–347 (2014).
26. Chi, D.-H., Kuo, W.: Optimal Design for Software Reliability and Development Cost. IEEE Journal on Selected Areas in Communications. 8, 276–282 (1990).
27. Shooman, M.L.: Reliability of Computer Systems and Networks – Fault Tolerance, Analysis, and Design. John Wiley & Sons New York, New York, NY, USA (2002).
28. Ardakan, M.A., Hamadani, A.Z.: Reliability–redundancy allocation problem with cold-standby redundancy strategy. Simulation Modelling Practice and Theory. 42, 107–118 (2014).
29. Chern, M.-S.: On the computational complexity of reliability redundancy allocation in a series system. Operations Research Letters. 11, 309–315 (1992).
30. Soltani, R.: Reliability optimization of binary state non-repairable systems: A state of the art survey. International Journal of Industrial Engineering Computations. 5, 339–364 (2014).
31. Meyer, J.F.: On evaluating the performability of degradable computing systems. IEEE Transactions on computers. 100, 720–731 (1980).
32. Ouzineb, M., Nourelfath, M., Gendreau, M.: Tabu search for the redundancy allocation problem of homogenous series–parallel multi-state systems. Reliability Engineering & System Safety. 93, 1257–1272 (2008).
33. Bosse, S., Splieth, M., Turowski, K.: Multi-Objective Optimization of IT Service Availability and Costs. Reliability Engineering & System Safety. 147, 142–155 (2016).
34. Greenberg, A., Hamilton, J., Maltz, D.A., Patel, P.: The Cost of a Cloud: Research Problems in Data Center Networks. SIGCOMM Comput. Commun. Rev. 39, 68–73 (2008).
35. Ciardo, G., Muppala, J.K., Trivedi, K.S.: SPNP: Stochastic Petri Net Package. In: Proceedings of the 3rd International Workshop PNPM. pp. 142–151. IEEE Computer Society (1989).
36. Kulturel-Konak, S., Smith, A.E., Coit, D.W.: Efficiently Solving the Redundancy Allocation Problem Using Tabu Search. IIE Transactions. 35, 515–526 (2003).
37. Walker, E.: The Real Cost of a CPU Hour. Computer. 42, 35–41 (2009).
38. Oppenheimer, D., Ganapathi, A., Patterson, D.A.: Why do Internet services fail, and what can be done about it? In: 4th Usenix Symposium on Internet Technologies and Systems (USITS) (2003).

How to Empower Users for Co-Creation – Conceptualizing an Engagement Platform for Benefits Realization

Martin Semmann[1] and Christian Grotherr[1]

[1] University of Hamburg, Department of Informatics, Hamburg, Germany
{martin.semmann, christian.grotherr}@uni-hamburg.de

Abstract. Organizations invest huge portions of their budget in IT with the goal to realize benefits as improving work practice and establishing new processes. To achieve this goal, users are engaged throughout projects by various methods and approaches. Nevertheless, after the completion of a project, users lack power and opportunities to further realize benefits and thus assuring the overall success of a project. To close this gap, we present the concept of an engagement platform that empowers users collectively to induce change initiatives that enhances the realization of benefits in the post-project phase. By doing so, benefits management practices undergo a paradigm shift from recent top-down management towards bottom-up realization of benefits. This change in perspective also incorporates a service systems perspective as it focusses on the dynamic configuration of actors and resources to enable value creation in a complex context.

Keywords: service system engineering, software introduction, technochange, user-generated services, benefits management

1 Introduction

Organizations invest huge portions of their budget in IT with the goal to realize benefits as improving work practice and establishing new processes [1, 2]. To achieve these objectives, IT investments must be well embedded in the organizational context resulting in complex project constellations. Additionally, anticipated benefits of the software can only be created in distinct contexts by various users utilizing the software. Thus, projects contribute to a service system, as a sociotechnical artifact in a distinct organizational environment is instantiated [3]. Following, benefits realization is done by using this sociotechnical artifact in a specific context while integrating various resources and actors [3]. Engaging users is therefore state of practice during projects by various methods and approaches [4, 5]. This engagement is done by selecting some users with a top-down approach within the project. This top-down approach is advantageous to get projects approved and delivered. Whereas a much broader or even general participation is complex, expensive and hard to keep target-oriented during a

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Semmann, M.; Grotherr, C. (2017): How to Empower Users for Co-Creation - Conceptualizing an Engagement Platform for Benefits Realization, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 91-105

project. Especially, considering major changes in software as introductions of new software or significant upgrades only representing users can be engaged efficiently throughout the project. Thus, most users cannot actively participate in the adaptation of software and organizational changes. Even more due to the context of use that is defined by the actors involved and the organizational boundaries this limited engagement leads to limited ability to realize benefits entirely. This limitation even increases after the completion of a project, users lack opportunities and power to further realize benefits and thus assuring the overall success of a project [6]. Recent literature reviews on benefits management from a project perspective [31, 32] show that, in post-project phase, there is no established method or concept to support emerging benefits as well as intended but unrealized benefits which is also reflected in a qualitative study [6]. This lack of engaging users is also mirrored as a third of installed software in organizations is estimated to be not used at all [7].

By utilizing a service systems perspective with the users as facilitators of value in context, a bottom-up approach seems more beneficial to enhance capturing of benefits to overcome these limitations in the post-project phase. Especially, regarding varying time lags and emergent benefits that have not been anticipated [4, 11-13]. Based on this perspective, a shift towards a bottom-up approach for enforcing co-creation within the community of users to further realize benefits and thus improving the solution and its value delivered collectively [8, 9]. A promising approach to instantiate such a bottom-up engagement platform is internal crowdsourcing as it aims for collaborative value facilitation within an organization by potentially engaging all users [33]. This active engagement also copes with the need for organizational change that complements new or changed IT to realize benefits [11]. This is also recognized in literature on IT-enabled transformation that emphasizes that capturing benefits is a critical post-project activity [10]. Following this argumentation, the paper answers the following research question: *How can a concept to empower users for co-creation of change initiatives be designed to enhance the possibilities to realize benefits?*

We do so by presenting the concept of an engagement platform that empowers users to collaboratively induce change initiatives that enhances the realization of benefits in the post-project phase. The resulting platform seeks to catalyze the potential of value co-creation as it decidedly addresses the context of users' engagement with the delivered software during the introduction. To enable value creation between actors of the service system, users should be empowered to implement change initiatives and thus, foster timely realization of benefits. This novel approach exceeds common crowd initiatives established for example within innovation management as change initiatives are not only identified and ranked, but explicitly realized within a specific organizational context.

Thus, benefits management practices undergo a paradigm shift from recent top-down management towards bottom-up realization of benefits. This shift has the potential to increase the ability to change organizations and their work practice drastically [14].

As service research [3] as well as design research [15, 16] calls for evidence-based cumulative research, we propose the concept to an engagement platform as the result of the design phase of our design science project. The remainder of the paper is therefore structured as follows: the second section builds up a foundation of the research by defining and summarizing related research. In the third chapter, we describe the methodology used to develop the engagement platform. All components of the concept are derived and comprehensively described in chapter four. The paper closes with a conclusion and outlines future research.

2 Conceptual Foundations

2.1 Service Systems Engineering

Service systems describe a configuration of actors and resources and their interaction [1] in order to enable co-creation of value by sharing resources among actors [2]. This is in line with the definition given by Böhmann et al. who conceptualize a service system as "complex socio-technical systems that enable value co-creation" [3]. Research has recognized the emergent importance of service systems and the need for establishing further research within this field such as service science [1, 4]. This research is supposed to address the interaction between actors regarding human agents with knowledge and skills as well as resources as technology, information, physical artifacts which interact in co-creation [1]. Service systems engineering elaborates therefore on the importance of systematic design and development of such service systems and calls for research on evidence-based design knowledge [3]. Service systems research consequently applies the principles of service-dominant logic which constitutes value creation through collaboration and contextualization [5]. Accordingly, contextualization emphasizes that producer and consumer create value collaboratively by configuring actors and resources specifically in a context [6, 7]. Hence, service systems enable value co-creation through configuration of actors and resources guided by its value proposition [5]. Understanding service systems as configuration of actors and resources with the aim of searching for principles and approaches that can help to improve value co-creation [8] we focus on the integration of these resources in order to foster the end-user co-creation of value within software implementation projects to realize benefits jointly.

2.2 Internal Crowdsourcing

Crowdsourcing is an IT-enabled phenomenon which is based on social IT like wikis, blogs or social networks [9]. Crowdsourcing can be defined as using information technology to connect various potential user groups to accomplished tasks by voluntary crowd workers often motivated by mutual benefits [10]. One main characteristic of crowdsourcing is the location of the crowd, which can be distinguished between external (e.g. communities of interest, customers) and internal (employees). External crowdsourcing has been applied in different industrial contexts as exemplified by the cases of LEGO [11] and SAP [12]. Yet, little is known about building and engaging a crowd within organizations [9]. As shown by Zuchowski et al., internal crowdsourcing has characteristics which distinguish it from external crowdsourcing. For example, the crowd is comprised of employees and is thus long-term oriented rather than independent ad-hoc and short-term-oriented external crowds [9]. An extensive literature review stated conflicting definitions and conceptualizations of internal crowdsourcing in literature [9]. The authors define internal crowdsourcing as “an (a) IT-enabled (b) group activity based on an (c) open call for participation (d) in an enterprise” [9]. This definition is in line with an engagement platform from a service systems perspective and therefore bears the potential to support benefits realization. Another characteristic is the need for organizational culture management skills, because the approach requires an open organization where employees can collaborate and debate with each other without having cultural boundaries [13]. A characteristic of external crowdsourced solutions, on the other hand, is that the design has the potential to reveal ‘outside the box’ information, while an internal crowd may also be suitable to solve contextualized, enterprise-centered problems [11]. In addition to location, the task is an important factor for distinguishing crowdsourcing approaches [14]. Crowds can be engaged to gain access to a diverse knowledge base as tasks vary between low levels of complexity, as considered in research on microtasking or microworking [15], to tasks with increasing complexity such as ranking, sharing knowledge, ideation to design and development of new solutions. While tasks with low complexity can be crowdsourced externally to increase productivity by reducing time and costs, knowledge-intensive tasks with a high complexity will often preferably be allocated to internal crowds as only an internal crowd is fully aware of a given context.

3 Research Design

The research project follows a design-oriented research strategy [16] and is conducted by utilizing the Design Science Research Methodology [17] to systematically and iteratively design, develop as well as demonstrate and evaluate a sociotechnical artifact in a suitable context.

Therefore, the first phase Problem Identification and Motivation aims for defining the research problem and adjusting the target of the solution. This deep understanding of the problem space defines the vision of the to be designed artifact. This research project follows the problem-centered initiation as the practical relevance is shown in the introductory section as well in following chapter. Although a lack of benefits realization targeted by software implementation projects is identified current research does not address this issue. This research therefore aims at developing a concept to empower users for co-creation of improvements to enhance benefits realization after software introductions.

In the following phase objectives of a to be designed solution are derived grounded on a previous study on post-project management in large organizations and research on service systems. The next phase Design and Development utilizes these results as the foundation of the implementation. As scholars call for cumulative research in service research [3] as well as design research [18, 19] we propose a concept as a result of the design and development phase as focus of this research. Nevertheless, as design, development, and demonstration are highly iterative phases, we include insights of the demonstration of early mock-ups and a first prototype that build the foundation of a future evaluation. This evaluation is planned to be guided by the Framework for Evaluation in Design Science (FEDS) [20]. Therefore, in the planned Evaluation phase the artifact is applied in the context of a Microsoft SharePoint introduction within the case organization. Thus, a suitable context to validate its applicability and utility by solving real problems is given [17]. The results gathered throughout this evaluation likely lead to further improvements on the initial concept.

4 Designing Benefit Realization Supporting Components

In the following section the course of the design science research project is described that leads to the design of the benefit-supporting components. The focus hereby lies on the conceptualization in the design and development phase. Accordingly, the first two phases are only shortly described as this project seeks for a cumulative communication of the results as called for by researchers [3, 18, 21].

4.1 Problem Identification and Motivation

Service systems have evolved into key concepts for research in information systems [1, 22]. Many industries such as IT manufacturing and healthcare seek to design effective technology enabled service systems that efficiently allow the configuration of the service system to meet individual needs and to create value in each context [3, 23]. As various studies show, a major problem of software introductions is that the resulting solutions is insufficiently used in organizations and thus, value is not created [24-27]. This lack of use varies from denial of use at all, users establishing workarounds to using a software but not efficiently or even effectively [25, 28, 29].

Despite this general problem description, this project is done in close cooperation with a client organization. The research takes place in a public law institution with 1.800 FTE. During an initiating workshop, the described problem was mirrored in this organization. Thus, a software introduction project was identified that fit to the described problem and has the potential to implement the to be designed concept of an engagement platform. Consequentially, the artifact aims at realizing benefits targeted by the project with a concept to empower users to co-create value within an engagement platform that integrates operant and operand resources within this service system. This is done by identifying possible improvements, discussing these, and applying the improvements collectively to realize benefits.

4.2 Objective of the Solution

With the overall problem definition as foundation for this design science research project, objectives of a solution must be identified. To do so, two approaches were taken. On the one hand, a preliminary qualitative study in twelve large organizations was conducted that evaluated the state of benefits management after a projects result is delivered [26]. The study reveals shortcomings of current practice that lead to implications for the design of the to be designed artifact (O1-4). On the other hand, literature on service systems engineering gives directions on the integration of resources and how actors can co-create value. Based on this research stream, a novel approach is taken that focusses on user-integration to co-create not only ideas for improving a software but also implementing the proposals by applying deep contextual understanding of engaging users (O5,6). The resulting objectives and their related sources are subsumed in Table 1.

Table 1. Objective of the proposed Solution

No.	*Objective*	*Source*
O1	Enforce continuity of benefits management that outlasts projects	[26]
O2	Accompany transition and early usage phases with an ongoing action-oriented approach instead of only a retrospective one	[26]
O3	Identify emergent benefits after the transition is completed and regular work practice is achieved	[26, 30]
O4	Establish ways to deal timely with improvements	[26, 31]
O5	Mobilize resources to enable user-driven change	[3, 32-35]
O6	Establish a platform that allows actors to engage	[33, 36]

The first objective considers the dynamic during projects and afterwards that ownership of benefits is changing dynamically (O1). Therefore, an engagement platform should ensure that change proposals are consistently related to the initiator or a governing actor to be able to take on actions that support progressing with the change. Thus, distinct actors are aware of the benefits related with the change and can monitor its realization. Additionally, they have the ability to communicate the usefulness. Secondly, practical insights show that current benefits management practice is mainly retrospective in the post-project phase. Therefore, a solution needs an action-oriented approach (O2) to enable actors to improve the deployed software according to the specific needs to ensure the realization of value in context. Hence, it is not sufficient to solely collect change requests to propose follow-up projects. As users establish work routines with the introduced software [37], a solution should support users by identifying further unintended benefits (O3). By doing so, users can be more engaged by improving the software and contextualize it based on their specific needs. Analogously, by establishing approaches to timely implement and thus improve the introduced software (O4) users' engagement is likely to increase and as a result benefits realization increases as well. As a major challenge in service systems engineering is the mobilization and integration of resources, a solution should incorporate approaches to do so (O5). Following Breidbach et al., the solution should have touch points that provide structural support for actors to realize the exchange and the integration of resources [36]. Finally, a solution to enable users to improve introduced software needs to be designed as an engagement platform (O6) [33, 36]. Consequently, the solution should facilitate exchange between users.

4.3 Design and Development

To address these objectives and as the third activity of the design science research process a concept is developed with the overall aim to enable end users to contribute to adaption and customization of an introduced software. Hence, the concepts integrate mechanisms to engage all users of a software recently introduced to exchange and integrate resources to improve the software. By striving for this goal a fundamental change takes place as an internal crowd is empowered to change software utilizing a bottom-up approach. This approach leads to empowered users that can propose, interact on, and realize changes to a software. In this context, opportunities are supported, which help to mobilize and access previously untapped resources of users leading to a contextualized adaptation of the software and thus bearing the potential to improve benefits realization [38]. Doing so facilitates and empowers users to build and strengthen capabilities for implementing change initiatives using dynamic resource integration as an internal crowd. This concept shifts benefits realization from strictly formalized processes towards support in collecting experience and perception of users directly affected using the new software.

As this research takes a problem-centered approach, the design is mainly driven by the aforementioned practical and theoretical insights. Due to the strong commitment of the client organization, each iteration that lead to this concept was demonstrated and refined with practitioners. Nevertheless, the concept represents an abstraction and therefore, can comprehensively be adapted to other contexts as well.

Following the objectives, the concept for empowering users to co-create change initiatives and to enhance benefits realization in software introductions consists of three core components. A user joins the engagement platform and follows the concept in a sequence by proposing a change initiative (C1). The second component (C2) aims for gaining crowd-commitment as supporting factor for realizing the change initiative and embody validation by the internal crowd if the change initiative is worthwhile realizing. Last, the third component (C3) supports users to realize change initiatives that are accepted by the crowd and deemed beneficial. However, the concept has an iterative character which allows re-entry in earlier components based on insights gained during the initial change initiative. Possible insights can be further change initiatives, spare change initiatives or insights which impacts the proposed change initiative.

Every component subsumes several functions that aim to transform an expected input into desired output. Subsequently, we describe the three core components of the concept in detail. We thereby focus on functions, their interfaces, cross-sectional dependencies, and design variables that need to be considered for instantiations of the concept in various service systems.

Proposing a Change Initiative (C1)

The aim of this component is to provide an engagement platform for users that enables them to collect ideas for change initiatives (Table 2). These initiatives are only emergent during the use of the introduced software in specific contexts. If for example, a process lacks accuracy during its runtime users can report immediately and contribute a change initiative for the redesign of this process. To propose a change initiative, users specify the change initiative (C1F1). This is done by describing the idea or issue (C1F2) and the related software as well as suggestions how a resolution could be realized on the engagement platform. To join the platform users should first create a user profile with information about skills and to further relate to matching change initiatives (C1F3). By using the platform, the profile will be extended with tags of interest for initiatives a user engaged with and thus represents a user's context holistically. Another mode to join the platform is to anonymously participate on the platform. This design decision must take into the effects of anonymity in communities' consideration as well as relatability of individual opinions. Table 2 subsumes the functions and highlights design decisions made in the organizational context of the project.

Table 2. Overview Component C1: Proposing Change Initiative

Objective O1, O2, O3	
Input idea statement, improvement proposal, solution design	
Functions	*Design Variables*
(C1F1) initialize change initiative	**idea, solution, problem**
(C1F2) describe change initiative	free text, **defined template**
(C1F3) create user profile	anonymous, **single-sign-on**, new profile
Output well formulated change initiative	

Gaining Crowd-Commitment (C2)

The overall aim of this module is to gain crowd-commitment for a proposed change initiative. Thus, users are supposed to engage to co-create suggestions and possible solution designs. Accordingly, one purpose of this component is to build communities of interests. To participate in such a community modes of crowdsourcing can be distinguished in general between the modes 'wisdom of the crowd' and 'marketplace/contest' [39]. With the aim of improving usage of software and with the boundary condition of limited members in the user base it is not suitable to compete against each other. Moreover, the overall aim is to work collaboratively on a solution to an identified problem. This is in line with the guiding definition of internal crowdsourcing which declare an 'open call for participation' [9]. Therefore, the concept should provide opportunities to discover change initiatives (C2F1). This can be instantiated using search and filter functions for new and relevant change initiatives. A more proactive and dynamic way to discover change initiatives is by demonstrating success stories related to user profiles by recommender engines.

Providing feedback for change initiative, developing suggestions and solutions (C2F2, C2F3) as well as rating change initiatives (C2F4) requires engagement between actors (C2F5). To prioritize change initiatives rating mechanisms can be implemented inspired by funding, voting and rating mechanisms. Based on the feedback and a prioritization change initiatives are selected which have particularly high and relevant benefits for software usage. To address a broad range of users, groups of interests and departments these functions must be provided across the organization to give all users the opportunity to participate as well as to involve users (C2F6). Therefore, communication such as blogs or forums are needed. Additionally, opportunities to address single users explicitly with sharing functions or with tagging systems that may suggest potential experts are needed to support communicating change initiatives and to engage users. A web-based information system which provides users a communication infrastructure is needed to allow them to share change initiatives, feedback, design discussions and helping to build solver groups. The participation of users will be strengthened in this way and they can contribute their expertise to provide improvements for a wider range of users. Gaining crowd-commitment does not only aim for gathering feedback for a change initiative but moreover to build a realization team to solve the issue and implement the developed solution design (C2F7). In this regard a user volunteers as a solver and thus teams up with the requestor and other committed users. This (virtual) formation can be supported for example by expertise matching tag systems as well as direct addressing potential solvers.

Table 3. Overview Component C2: Gaining Crowd-Commitment

Objective	O2, O4, O5, O6
Input	change initiative
Functions	*Design Variables*
(C2F1) discovering change initiative	**search function**, **success stories**, recommendations, **filter function**
(C2F2) feedback change initiative	**blog**, forum, instant messaging
(C2F3) develop suggestions and solutions	**free text**, mock-ups
(C2F4) rate change initiative	funding, rating, **voting**
(C2F5) communicating change initiative	**passive**, **active**
(C2F6) involve users, experts	**tagging**, mail, newsletter
(C2F7) building solver-team	**self-organized**, direct communication
(C2F8) govern crowd	**self-regulating**, passive controlling, community-manager
(C2F9) monitoring status change initiative	promote, remove, **provide status**
Output	(virtual) team formation, refined and validated solution design

Further mechanisms should be considered that adopt functions of managing the crowd. For example, in the case of inadequate comments guidance how to govern the crowd are required (C2F8). This might imply the need for community management as well as reporting mechanism. Additionally, by monitoring the status of a change initiative and information about recent activities, community management can actively promote or remove outdated change initiatives (C2F9). The hurdle lies in the activation of users to engage on the platform, discovering change initiatives and to participate with feedback, rating as well as solving change initiatives. Guided by the demand to design an “engagement platform to incentivize certain actors to contribute their resources and enable service-for-service exchange” [33], corresponding motivation, activation and incentive mechanism for users have to be established. Therefore, motivation and incentives can be distinguished between the source of incentive (intrinsic, extrinsic) and the object (monetary, non-monetary) [40] and should be embedded in the instantiation of the concept [41]. However, the willingness and openness to participate on the engagement platform may be restricted by social influences. By designing communication, coordination, motivation and incentive guidelines the boundaries of individual decision making within an organization and closed communities should be considered. Actors act within a structure restricted by social rules and collective meanings, which are part of the organizational culture [42]. This is mirrored as well in the overview given in Table 3 including the design decisions in the case organization.

Realizing Change (C3)

As the overall aim of the concept is to realize change initiatives. As organizational context also embodies limited time for additional activities and lack of access permissions, change initiatives will be implemented jointly by the crowd and transferred to regular operation (C3F1). By providing dedicated time for users or adding additional resources users are empowered to realize benefits for themselves and for other users (C3F2). It is also possible that projects arise, which are equipped additionally with budgets and possibly additional resources and handed over to general project management. Other ways to support realization of change initiatives are crowd mechanism (C3F3) such as task management [43]. Building tasks to split workload and provide the possibility for lightweight participation in the realization process. Further dividing realization projects into small tasks supports automated testing and automatic integration [43]. After users have realized a change initiative, the solution should be tested and evaluated regarding defined acceptance criteria (C3F4). This also depends on the context and thus needs to be defined during instantiation of the engagement platform. After realizing and deploying change initiatives engaged users are informed and rewarded as defined during instantiation of the engagement platform (C3F5).

Table 4. Overview Component C3: Realizing Change

Objective	O2, O4, O5
Input	solution design
Functions	*Design Variables*
(C3F1) realizing change initiative	
(C3F2) enable realization	**attracting experts/consultants/IT**, providing dedicated time
(C3F3) building, assigning tasks	**self-regulated**, supported by tools, only if no additional tools are needed
(C3F4) testing and evaluating change initiative	how (**not mandatory**, acceptance criteria), who (IT department, **user**)
(C3F5) reward participants	monetary, **non-monetary**
Output	realized, deployed change initiative, realized benefits

4.1 Demonstration of a Preliminary Instantiation

The conceptual results of each design and development cycle were already initiated as prototypes and demonstrated within the case company. Starting with a reduced prototype the demonstration of the components and their functionality was initially conducted with a low-fidelity prototype (mock-ups). By extending the concept incrementally based on the preliminary results of the demonstration, the overall concept was instantiated as a responsive web application based on open source frameworks as shown in Figure 1.

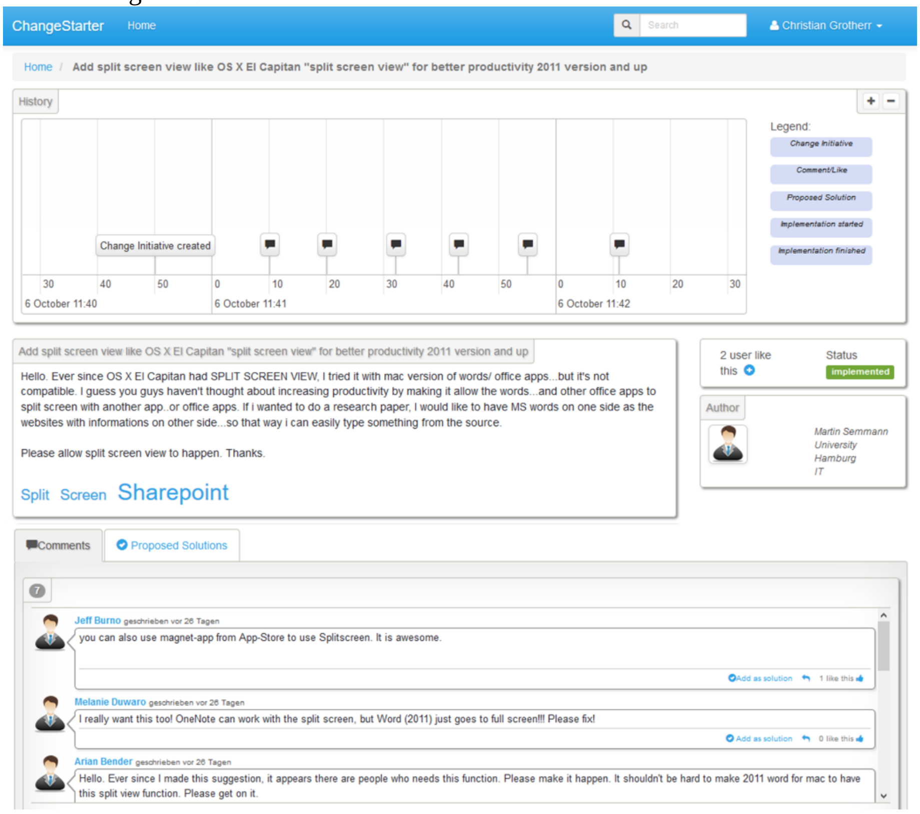


Figure 1. Instantiated User Engagement Platform supporting Benefits Realization

In sum, five workshops were conducted lasting two hours each including highly relevant stakeholders such as the CIO, head of IT operations, senior managers, representatives of the workers' council, and privacy commissioner to gain strong commitment of management as well as workforce.

Within the demonstration phase, feedback was gathered regarding the set of design variables and their manifestation to meet the requirements of the organization like the condition of voluntary and autonomous participation on the engagement platform. The results are highlighted in Table 2 to 4. Additionally, further extensions and improvements of features were discussed. For example, features were added to support discovering change initiatives (C2F1) like search functions and success stories. Despite this, every workshop helped streamlining the overall usability by simplifying the user interface to decrease adaption barriers.

4.2 Evaluation

As the first completed demonstration of the concept and its instantiation was successful, an extensive evaluation is currently planned. This evaluation is will be operationalized at the case organization and is open to all employees. Based on the gained commitment of relevant stakeholders during demonstration phase, we can deploy the prototype within the systems of the client and ensure deliberately low participation. Moreover, the evaluation does not have a dedicated timeframe and thus the internal crowd of the organization can evolve over time. The goal is to include 100 FTEs during the first phase of the evaluation. To achieve this goal, a set of potentially interested users is identified that could act as promotors for the concept within the organization. These users also serve as pre-tester to populate the platform with initial initiatives.

By evaluating the artifact within the organization, feedback is gathered applying qualitative methods such as interviews or thinking aloud to get insights on user's perception [44, 45] as well as gathering usage data. Accordingly, we do not only focus on the technical evaluation but also seek to gain insights on the social consequences of the artifact. Thus, the evaluation will contribute to the ongoing debate on socio-technical artifacts [46, 47]. The experiences and results of the evaluation are directly incorporated into further development and refinement of the concept.

5 Conclusion

Striving for a rise of benefits realization after a software introduction is formally closed, we presented a novel concept of an engagement platform. This concept utilizes a service systems perspective to empower users by a bottom-up approach to propose, engage and discuss and finally implement changes for this software and work routines. By doing so, the entirety of users can improve sociotechnical interaction to enhance the creation of value in context. Consequently, users are empowered to realize benefits that could not sufficiently be addressed during the software introduction project but even more, can deal with emergent benefits collectively. As the design of the concept integrates practice-oriented as well as theoretical insights within a case organization to instantiate the concept, in depth knowledge on the integration of resources in a complex service system as well as engagement strategies can be gained. Thus, this research is a core foundation towards an evaluation that is evidence-based and bears the potential to further improve design knowledge on actor-centered service systems engineering. Additionally, the proposed concept relates to current research on benefits management that seeks to understand how benefits realization can be fostered on actor level.

As a next step, the concept will be evaluated in practice within the introduction of Microsoft SharePoint. Moreover, it is planned to apply the concept to other contexts to assess and further enhance the transferability. Especially, regarding the design variables we seek to identify beneficial combinations to strengthen the engagement of users and thus contribute to the still emerging research on actor engagement in service systems.

Acknowledgement

This research was partly sponsored by the German Federal Ministry for Education and Research in the project ExTEND under the reference 01FJ15125. Further information can be found under: http://www.projekt-extend.de/. Additionally, we thank Philipp Streicher for his efforts to prototypically implement the platform.

References

1. Alter, S.: Metamodel for service analysis and design based on an operational view of service and service systems. Service Science 4, 218-235 (2012)
2. Maglio, P., Vargo, S., Caswell, N., Spohrer, J.: The service system is the basic abstraction of service science. Information Systems and e-Business Management 7, 395-406 (2009)
3. Böhmann, T., Leimeister, J., Möslein, K.: Service Systems Engineering. Business & Information Systems Engineering 6, 73-79 (2014)
4. Maglio, P.P., Spohrer, J.: Fundamentals of service science. Journal of the Academy of Marketing Science 36, 18-20 (2008)
5. Vargo, S.L., Lusch, R.F.: Evolving to a new dominant logic for marketing. Journal of Marketing 68, 1-17 (2004)
6. Edvardsson, B., Tronvoll, B., Gruber, T.: Expanding understanding of service exchange and value co-creation: a social construction approach. Journal of the Academy of Marketing Science 39, 327-339 (2011)
7. Vargo, S.L., Maglio, P.P., Akaka, M.A.: On value and value co-creation: A service systems and service logic perspective. European Management Journal 26, 145-152 (2008)
8. Vargo, S.L., Akaka, M.A.: Value cocreation and service systems (re) formation: A service ecosystems view. Service Science 4, 207-217 (2012)
9. Zuchowski, O., Posegga, O., Schlagwein, D., Fischbach, K.: Internal Crowdsourcing: Conceptual Framework, Structured Review and Research Agenda. Journal of Information Technology (forthcoming)
10. Estellés-Arolas, E., González-Ladrón-De-Guevara, F.: Towards an integrated crowdsourcing definition. Journal of Information Science 38, 189-200 (2012)
11. Schlagwein, D., Bjørn-Andersen, N.: Organizational learning with crowdsourcing: The revelatory case of LEGO. Journal of the Association for Information Systems 15, 754 (2014)
12. Leimeister, J.M., Huber, M., Bretschneider, U., Krcmar, H.: Leveraging crowdsourcing: activation-supporting components for IT-based ideas competition. Journal of Management Information Systems 26, 197-224 (2009)
13. Benbya, H., Van Alstyne, M.W.: How to find answers within your company. Sloan Management Review 52, 66-77 (2010)
14. Erickson, L.B.: Leveraging the crowd as a source of innovation: does crowdsourcing represent a new model for product and service innovation? In: Proceedings of the 50th annual conference on Computers and People Research, pp. 91-96. ACM, (2012)
15. Brabham, D.C.: Crowdsourcing. MIT Press (2013)
16. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS quarterly 28, 75-105 (2004)

17. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. Journal of Management Information Systems 24, 45-77 (2007)
18. Niederman, F., March, S.T.: Design science and the accumulation of knowledge in the information systems discipline. ACM Transactions on Management Information Systems (TMIS) 3, 1 (2012)
19. Iivari, J.: Distinguishing and contrasting two strategies for design science research. European Journal of Information Systems 24, 107-115 (2015)
20. Venable, J., Pries-Heje, J., Baskerville, R.: FEDS: a Framework for Evaluation in Design Science Research. European Journal of Information Systems 25, 77-89 (2016)
21. Iivari, J.: Distinguishing and contrasting two strategies for design science research.
22. European Journal of Information Systems 24, 107-115 (2015)
23. Fielt, E., Böhmann, T., Korthaus, A., Conger, S., Gable, G.: Editorial: Service Management and Engineering in Information Systems Research. The Journal of Strategic Information Systems 22, 46-50 (2013)
24. Ostrom, A.L., Parasuraman, A., Bowen, D.E., Patricio, L., Voss, C.A., Lemon, K.: Service research priorities in a rapidly changing context. Journal of Service Research 18, 127-159 (2015)
25. Ward, J., De Hertogh, S., Viaene, S.: Managing benefits from IS/IT investments: An empirical investigation into current practice. In: System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on, pp. 206a-206a. IEEE, (2007)
26. Limited, E.: The Real Cost of Unused Software. In: Limited, E. (ed.), (2015)
27. Semmann, M., Böhmann, T.: Post-Project Benefits Management in Large Organizations – Insights of a Qualitative Study. In: International Conference on Information Systems (ICIS), pp. 16. (2015)
28. Marchand, D.A., Kettinger, W.J., Rollins, J.D.: Information orientation: people, technology and the bottom line. Sloan Management Review 41, 69-80 (2000)
29. Roder, N., Wiesche, M., Schermann, M., Krcmar, H.: Toward an Ontology of Workarounds: A Literature Review on Existing Concepts. In: 49th Hawaii International Conference on System Sciences (HICSS), pp. 5177-5186. IEEE, (2016)
30. Zainuddin, E., Staples, S.: Developing a Shared Taxonomy of Workaround Behaviors for the Information Systems Field. In: 49th Hawaii International Conference on System Sciences (HICSS), pp. 5278-5287. IEEE, (2016)
31. Majchrzak, A., Markus, M.L., Wareham, J.: Designing for digital transformation : lessons for information systems research from the study of ICT and societal challenges. MIS Quarterly 40, 267-277 (2016)
32. Patora-Wysocka, Z.: The institutionalization of practice: a processual perspective on value co-creation 1. Economics and Business Review 2, 113 (2016)
33. Lusch, R.F., Vargo, S.L., Gustafsson, A.: Fostering a trans-disciplinary perspectives of service ecosystems. Journal of Business Research 69, 2957-2963 (2016)
34. Storbacka, K., Brodie, R.J., Böhmann, T., Maglio, P.P., Nenonen, S.: Actor engagement as a microfoundation for value co-creation. Journal of Business Research 69, 3008-3017 (2016)
35. Vargo, S.L., Lusch, R.F.: Institutions and axioms: an extension and update of service-dominant logic. Journal of the Academy of Marketing Science 44, 5-23 (2016)
36. Peters, L.D., Löbler, H., Brodie, R.J., Breidbach, C.F., Hollebeek, L.D., Smith, S.D., Sörhammar, D., Varey, R.J.: Theorizing about resource integration through service-dominant logic. Marketing Theory 14, 249-268 (2014)
37. Breidbach, C., Brodie, R., Hollebeek, L.: Beyond virtuality: from engagement platforms to engagement ecosystems. Managing Service Quality 24, 592-611 (2014)

38. Bapuji, H., Hora, M., Saeed, A.M.: Intentions, intermediaries, and interaction: Examining the emergence of routines. Journal of Management Studies 49, 1586-1607 (2012)
39. Breidbach, C.F., Maglio, P.P.: A service science perspective on the role of ICT in service innovation. ECIS 2015 Research-in-Progress Papers. (2015)
40. Vukovic, M., Bartolini, C.: Towards a Research Agenda for Enterprise Crowdsourcing. In: Margaria, T., Steffen, B. (eds.) Leveraging Applications of Formal Methods, Verification, and Validation, vol. 6415, pp. 425-434. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
41. Przygodda, I.: Anreizsysteme zur Bildung und Steigerung der Motivation für den Wissenstransfer. In: Zelewski, S., Ahlert, D., Kenning, P., Schütte, R. (eds.) Wissensmanagement in Dienstleistungsnetzwerken: Wissenstransfer fördern mit der Relationship Management Balanced Scorecard, pp. 59-95. Deutscher Universitätsverlag, Wiesbaden (2005)
42. Cuel, R., Morozova, O., Rohde, M., Simperl, E., Siorpaes, K., Tokarchuk, O., Wiedenhoefer, T., Yetim, F., Zamarian, M.: Motivation mechanisms for participation in human-driven semantic content creation. International Journal of Knowledge Engineering and Data Mining 1, 331-349 (2011)
43. Lusch, R.F., Nambisan, S.: Service Innovation: A Service-Dominant Logic Perspective. MIS Quarterly 39, 155-175 (2015)
44. Dwarakanath, A., Chintala, U., Shrikanth, N.C., Virdi, G., Kass, A., Chandran, A., Sengupta, S., Paul, S.: Crowd Build: A Methodology for Enterprise Software Development Using Crowdsourcing. In: 2nd International Workshop on CrowdSourcing in Software Engineering (CSI-SE), pp. 8-14. IEEE/ACM, (2015)
45. Boren, T., Ramey, J.: Thinking aloud: Reconciling theory and practice. IEEE transactions on professional communication 43, 261-278 (2000)
46. Myers, M.D.: Qualitative research in business and management. Sage (2013)
47. Silver, M.S., Markus, M.L.: Conceptualizing the SocioTechnical (ST) artifact. Systems, Signs & Actions 7, 82-89 (2013)
48. Goldkuhl, G.: From ensemble view to ensemble artefact–an inquiry on conceptualisations of the IT artefact. Systems, Signs & Actions 7, 49-72 (2013)

Design and Evaluation of a Smart-Glasses-based Service Support System

Christina Niemöller[1], Dirk Metzger[1] and Oliver Thomas[1]

[1] Universität Osnabrück, Informationsmanagement und Wirtschaftsinformatik, Osnabrück, Germany
{christina.niemoeller,dirk.metzger,oliver.thomas}@uni-osnabrueck.de

Abstract. The character of IT transformed from an attached commodity to the center of new products and services. Especially in technical customer services, new technologies such as smart glasses offer great opportunities to overcome current challenges. Due to the complexity of service systems engineering, guidance on how to design smart glasses-based service support systems is necessary. To overcome this complexity and fill the research gap of design knowledge, we (1) analyze the domain in a multi-method approach eliciting meta-requirements, (2) propose design principles, and (3) instantiate them in a prototype. We follow a design science research approach combing the build-phase with four evaluation cycles obtaining focus groups twice, demonstration with prototype and, based on that, a survey with 105 experts from the agricultural sector. We address real-world problems of information provisioning at the point of service and, thereby, contribute to the methodological knowledge base of IS Design and Service Systems Engineering.

Keywords: Service Systems Engineering, Service Support Systems, Smart Glasses, Design Science Research.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Niemöller, C.; Metzger, D.; Thomas, O. (2017): Design and Evaluation of a Smart-Glasses-based Service Support System, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 106-120

1 Introduction

Over the last decade, there has been an increasing focus on service science coupled with the design of new information and communication technologies [1, 2]. Especially concerning technical customer services (TCS) and their inherent field service, the support by mobile devices is inevitable. Service support systems are needed in order to cope with the wide range of service tasks [3, 4]. This can only be achieved through sufficient support by IT that proactively provide information and empower the service technician [5–7]. Although the (ergonomic and time-saving) potential induced by traditional head-worn-displays has already been discussed in research in the field of maintenance [8, 9], still only few scientific papers exist that address the emerging technology "smart glasses", a mobile eyewear with a display and features such as camera or sensors e.g. for head movement [10] as for example the Google Glass or the Vuzix M100. The existing papers mostly focus on individual and detached aspects, such as technical issues or scenarios like barcode scans [e.g. 11-13], none of them determine the design of a whole system.

We argue that there is a special need to examine the service support based on smart glasses. Based on a systematic literature study, Herterich et al. [14] identified the future research need of analyzing which field service tasks could be supported by innovative mobile technology. In particular the features of smart glasses offer new opportunities and enable the support of service technician (e.g. when free or clean hands are mandatory); thus, a study with focus on the device itself suit recent needs [10]. Nevertheless, to date, little research provides guidance for researchers and practitioners on how to build a smart glasses-based service support system. Against this background, the questions that guide our research are: (1) *What are the meta-requirements for a system that supports TCS in a hands-free way?,* (2) *How should smart glasses-based service support systems be designed that addresses these requirements?* (3) *How does the addressed user group evaluate the system regarding their intention to use it?*.

We followed a design science research approach (DSR) after Hevner et al. [15]. We conducted four evaluation cycles according to the Human-Risk & Effectiveness-oriented evaluation strategy proposed by Venable et al. [20]. By answering the proposed research question and presenting design principles as well as an evaluated instantiation as research artefact [16, 17], we contribute to the knowledge base of IS Design and Service Systems Engineering (SSE). With the several evaluation cycles including experts from theory and practice as well as researching in a real-world scenario, we address the call of SSE for research on evidence-based design knowledge for systems that permeate our society [18]. The derived design knowledge also guide practitioners by implementing a mobile service support system and enables them to create new business models (e.g. customer self-service).

We proceed as follows: First, we introduce the theoretical foundations of mobile service support systems for TCS. In section 3, we introduce our research approach. Next, we present the artefact design comprising the meta-requirements, design principles and the instantiation. In section 5, the results of the final evaluation are presented. We conclude by discussing novelty, practical relevance, theoretical contributions, and limitation as well as giving an outlook for future work.

2 Related Work

For many manufacturers technical customer services (TCS) became a major value-adding resource [4, 19]. In order to assist the service technicians, more and more researchers claim for the need of mobile service support systems [5–7, 20].

Due to the high range of tasks [3, 4] combined with the increasing complexity of high-tech products being subject to their work [21], TCS processes are complex entities. A service process of the TCS involves activities undertaken to realize and deliver the service at the so called point of service [5]. For his work on site he or she is dependent on current information about the whole service process. Conducting a case study, the authors Becker et al. [22] have analyzed information needs within service and manufacturing business processes of a milling/turning machine producer. They focus on how an integration of services and manufacturing can be accomplished by sharing information in service systems. In addition to that, Däuble et al. [23] derived information needs from literature and evidenced their investigation by results of real-world service process observations in the field of the machinery and plant engineering for the intralogistics sector. Thereby, they elicited 13 information needs such as information from the manufacturer, service item information, procedure information or tool information. In line with the authors, we focus on the information provision on site with essential information to fulfill the service tasks.

Agnihothri et al. [20], Legner et al. [7], and Ray et al. [6] focus on the impact of technology use on service performance while Fellmann [24] focus the proper integration of existing IS. Since diverse requirements have to be considered spanning technical aspects like interfaces or integration technology and the functionality of such systems, the development of a service support system is a complex task. In order to respond to this complexity and give guidance for further researcher, Matijacic et al. [5] elicited and consolidated requirements and mapped them to an generalized service process. So, as to embrace this richness, the authors suggest to use three distinct methods to elicit and consolidate requirements relying on different sources (systematic literature study, observations, and expert interviews). Likewise, we used triangulation for gaining the requirements for our smart glasses-based service support system. To guarantee the quality of the service, it is important to process and structure existing data to support the operation and the staff efficiently [6]. Considering the data and information flow between the system and the service technician, a bidirectional channel has to be established. Information is not only provided by the system, generated information and data while executing services are also carried back to it [5].

Additionally, based on a systematic literature study, Herterich et al. [14] identified the future research need of analyzing which field service tasks could be supported by innovative mobile technology such as wearables. This is where Niemöller et al. [10] start. They examine the features of smart glasses and map them to process steps of the TCS to show its potential for service support. We base on this work, when talking about the features such as hands-free interaction with e.g. voice control. Metzger et al. [25] propose to use smart glasses in the context of TCS to model service process during service provision. They also suggest to use smart glasses for a hands-free support during service work; however, they do not state how the system has to be designed.

3 Research Approach

We follow a classical design science research approach (DSR) [15, 26] as it is generally accepted for Service Systems Engineering (SSE) [18]. Böhmann et al. [18] propose that research needs to be embedded within a service system in a real-world scenario and call for the design of novel service systems. In line with the authors, our approach continuously involves experts from TCS as well as observations of real-world process scenarios. Following DSR, we investigated the four phases analysis, design, evaluation and diffusion as shown in Figure 1.

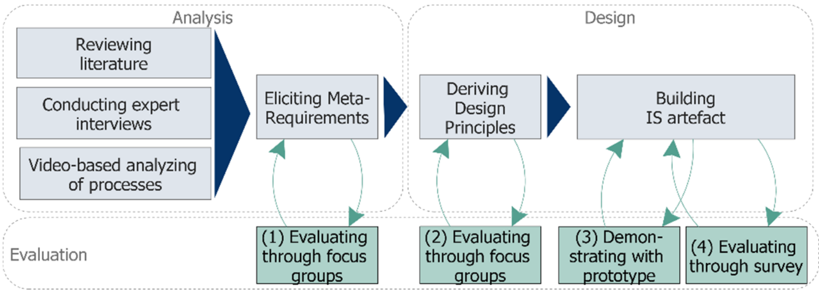


Figure 1. Design science-oriented research approach

Following the argumentations of practitioners (after discussions with several experts from the TCS in the agricultural industry as well as the air-conditioning sector) and service researcher, a need for support TCS trough IT was combined with a hands-free system interaction (cf. 1 and 2). Once the business problem is identified, attributes of the pursued future system have to be investigated and defined [15]. These attributes are usually referred to as *meta-requirements* [17] (cf. 4.1). The meta-requirements were elicited from the analysis of the real-world scenario (process analysis, expert interviews) and the IS knowledge base (systematic literature study). We asked two companies from agricultural and air-conditioning technology to use action cams or smart glasses to capture videos of how they execute service processes. Overall, 10 videos were captured showing different maintenance processes. We chose to conduct a triangulation to combine the different points of views and calibrate and validate our work [27]. The elicited requirements were mapped to the TCS process phase in which they appeared (order preparation, execution, post processing and phase-independent). We received a consolidated list and ranked the requirements by their significance within every process phase (analogously to Matijacic [5]). The highest rated requirements were discussed with the focus group. Next, an information system needs to be designed that meets the identified meta-requirements [17]. We defined the *design principles* (DP) based on the derived meta-requirements and literature, combining wearable computing design (e.g. [28–30]) with implications from service systems engineering and business process management (e.g. [3, 10, 21, 30]). We finalized the DPs catalogue after a second workshop with the same participants as in the first focus group meeting (for evaluating the meta-requirements), when all stakeholders were satisfied (cf. 4.2). Finally, the IT artefact was instantiated (cf. 4.3).

Since the evaluation of design artefacts and design theories is a central and critical part of DSR [15, 31], we combined the build-phase with several evaluation phases. Hence, Venable et al. [32] propose a framework for developing an appropriate evaluation strategy. Following their argumentation, our evaluation strategy is Human-Risk & Effectiveness-oriented. As a result, we have to evaluate our artifact and the design decisions early in a naturalistic setting, conducting formative and summative evaluations [32]. For implementing the evaluation strategy and choosing suitable evaluation methods, we made use of the proposed principles by Sonnenberg and vom Brocke [33]. According to the authors, we conducted four evaluation steps (cf. Figure 1): (1) Within the first evaluation (focus group), we verified whether the research need is important and novel to address a justified research gap. (1) Representing the TCS perspective from practice, three attendees from a small and medium-sized service provider for air-conditioning technology and three participants from large agricultural technology manufacturer with own TCS attended. (2) For gaining insides from a technological perspective, two IT practitioners and two visual technology researcher participated. (3) To bridge the technological and service view, three IS researcher specialized in service science were invited and took up the role as leader of the open discussion. (4) For the design of the content and targeted communication of information, two researchers with specialty in education and media psychology were invited. We received justified design objectives in form of verified meta-requirements. (2) Within the second evaluation (same focus group), we examined the feasibility, clarity, internal consistency and applicability of our DPs to gain a validated design specification. (3) By demonstrating our IS instantiation with a prototype, we proofed the feasibility as well as the suitability while discussing the demonstration with a selected group of expert from research and practice (same focus group plus five service technicians from the mentioned sectors). (4) The ex-ante evaluation cycles informed our work, e.g. that the acceptance of smart glasses plays a major role in the view of practitioners. This is why we demonstrated and evaluated the system on the world's biggest agriculture technology fair; and hence, validated the applicability to real world problems, the generality concerning different user groups as well as the ease of use. First, we explained the system functionalities (example case see Figure 2) and the interaction with the smart glasses (hand and voice recognition) to each individual participant while the participant was wearing the smart glasses her-/himself. After that, every participant could test the smart glasses system individually. After each demonstration, we asked the experts (n=105) about the acceptance by conducting a survey based on the Technology Acceptance Model (TAM) [34, 35].

4 Artefact Design

Meta-Requirements. Through the mixed-method approach we generated an overall of seven meta-requirements (MR). All of them were discussed with our focus group on a workshop. Table 1 describes the MR and their origin from the triangulation.
MR1: Process information. The requirement that was mentioned the most in the conducted interviews and was needed in every step within the video analysis was about

guiding the technician through the process. To ensure quality, some companies are using electronic or paper-based checklists to remind the technician of the most important steps. So, one advancement of a system that helps the technicians in the field is to provide checklists or a step-by-step guidance. Thus, it is raised as our basic MR. MR1 is in line with the information need N5, proposed by Däuble et al. [23].
MR2: Additional information. Besides the step-by-step guidance, we derived the need for additional information attached to a single step from interviews and video analysis. Reasons mentioned for that were: The technician (1) has never or rarely done that particular action before, (2) needs some details about the tools that have to be used for that step, (3) need information about the machine itself, such as technical details or spare part information. MR2 is in line with the information needs N1, N7, and N8 [23].

Table 1. Meta-requirements derived from triangulation

	Meta-requirements	*Interview*	*Process*	*Literature*
Functional	MR1: Process information.	X	X	[23]
	MR2: Additional information.	X	X	[23]
	MR3: Order overview.	X	X	[23]
	MR4: Order details.	X		[23]
	MR5: Feedback integration.			[6] , [5], [25]
Non-Functional	MR6: Hands-free interaction.		X	[14], [10]
	MR7: Usability.	X		

MR3: Order overview. Mentioned by one interviewee and also found in the video analysis, the need for an overview of the orders that the technician has to fulfill is given. When starting to work in the morning, an overview of the orders helps the technician to estimate how much time is calculated per order and how much work has to be done during the day. MR3 is in line with N10 [23].
MR4: Order details. Besides the overview of orders, we found multiple information that is attached to an order (kind of order and who issued it, related machine, machine and service history, maintenance contract) that needs to be provided to the technician. Within the interviews, order details were rated positive. MR4 is in line with N2, N3, N4, and N12 [23].
MR5: Feedback integration. When we evaluated the other six MRs, one challenge arose considering the document basis (the documented service processes) itself, as the service documentations need to be maintained and updated regularly. Within the discussion we found, that the one who know first, if there is something wrong with the processes, is the technician in the field. So, one crucial factor for the system is to include feedback of the technicians (to give advice that there is something missing, wrong or outdated). MR5 is in line with the argumentation by Ray et al. [6] and Matijacic et al. [5] about the bidirectional channel between technicians and administration.
MR6: Hands-free interaction. During our analysis of the technician's work, we asked them to capture a video of the maintenance process for multiple reasons. One of them was to analyze how often they are using one or two hands during the process. With the

two-hand ratio of about 80% (almost 6 of 7.5 minutes) we validated the fact that for most of the steps the technicians need both hands. This lead to our additional MR of a hands-free interaction with the system to ensure that the technician can continue his work while being assisted. MR6 is in line with the claim by Herterich et al. [14] for investigation of wearables in TCS and the analysis by Niemöller et al. [10].

MR7: Usability. During interviews and the workshop, one of the most important concerns about a support system was, that the technicians are distracted or overwhelmed by the system and are not willing to use it. So, one main aspect of the system is to be integrated into the work environment of the technicians with respect to easy and efficient usage. We added this as one MR, although it is not a requirement on functionality of the system but rather a requirement on how every aspect of the system needs to be designed. MR7 was emphasized during the focus group discussion.

Design Principles. We defined design principles (DP) to support the design of the system based on the derived MRs and literature, combining wearable computing design (e.g. [28–30]) with implications from service systems engineering and business process management (e.g. [3, 10, 21, 30]). The seven DPs are described in the following, starting with the ones that meet the non-functional MRs. If not state explicitly, the principles are generic for all smart glasses applications.

DP1: Use voice recognition of smart glasses as main interaction pattern. As the system should be usable during service delivery, a solution that makes sure the hands are free and usable is needed; hence, interaction based on hands such us buttons or gestures are inappropriate and should be complemented with sensor based interaction (following the first principle of [28]). However, the technicians need to interact with the system; for example, to update their progress or bring additional information to the front. After having a look at different interaction approaches, the least disturbing and most versatile interaction pattern is the usage of voice recognition [10]. So, the first DP is the usage of voice recognition as main interaction pattern in order to fulfill the hands-free interaction MR (MR6) and, thereby, generate additional value compared to other devices (as proposed in principle 4 by [28]).

DP2: Keep the menu navigation depth as small as possible. The technician need to find orientation and interact with the system in a very short amount of time to use the system efficiently [5, 21]. Complex software often involves complex menu navigation to enable the adjustment of all details. However, in combination with smart glasses the menu navigation is limited due to small screen area. Additionally, the system is supposed to be used by technicians during work; so, the main mental focus of the technicians is on the service delivery. Thus, our DP is to limit menu navigation depth to keep the interaction simple (in line with principle 3 in [29] and principle 5 of [28]). This contributes to the easiness of use of the software (MR7).

DP3: Always return to the last shown step. Considering the characteristics of TCS processes (complex and branched e.g. due to comprehensive fault detection trees) [3], the technician have to proactively get to the correct step without manual search [21]. So, the system needs to make sure that the progress for every order is saved and loaded correctly. Furthermore, when additional information is displayed or the user is giving feedback, the system needs to make sure that it is returning to the correct step in the process. Together with the last DPs this principle also improves efficiency and easiness to use (MR7). This DP was added after discussion with the focus group (evaluation 2). It is a specific DP for step-by-step guidance related activities.

DP4: Build an order management. Based on the need for an order overview (MR3) and order details (MR4) (which are specific MRs for order related activities), we included this principle. To overcome the missing screen area, the DP contains to build one screen for each day (overview) and one for each order (details). As smart glasses limit the amount of information that can be displayed at once, we propose to separate the details for every order from the overview (following principle 5 of [28]).

DP5: Build one main screen with crucial information about the step. The main functionality of the system is to give step-by-step guidance through the process. The key information for every step needs one screen that is easy to recognize and to understand. However, all important information needs to be included which brings the screen design in conflict between readability and completeness of information. The system designer needs to be aware of this interplay. Overall, when designed correctly, this principle contributes to the MRs step-by-step guidance (MR1) as well as usability (MR7) (following principle 5 of [28]). This is a specific DP for step-by-step guidance related activities.

DP6: Attach additional information such as texts, pictures and videos to specific steps. Every additional information ranging from spare part information, pictures, wiring diagrams, videos, technical details etc. needs to be included into the step-by-step guidance. We propose to attach the information in the data storage directly to the step where it might be needed. The relation between additional information and step might be implemented as many-to-many-relation (m:n) as there might be additional information that assists in multiple steps as well as multiple additional information for one step. With the relation, the additional information would be accessible when needed and makes sure long search periods are unnecessary (in line with [30]). This supports MRs additional information (MR2) as well as usability (MR7) due to easy access to additional information. This is as well a DP for step-by-step guidance related activities.

DP7: Allow direct feedback to one step. Finally, the integration of feedback functionality is essential to ensure data quality of the processes and fast alteration when needed. To design the feedback as easy as possible, we propose to make the feedback functionality accessible directly from the step. Consequently, when sending feedback, the information about the context such as the order information, customer information, information about the step where the feedback was sent etc. needs to be logged and included. This enables the administrator of the processes to assess the context of the problem and adopt the process accordingly (e.g. by changing the process for a special kind of customer). Thus, this contributes to "the fit between business processes and technology" [36] as it enables continuous adaptation. Overall, the integration of feedback contributes to the MR give feedback about content and processes (MR5) as well as usability (MR7) because of direct communication with the process administration. Finally, this is a specific DP for step-by-step guidance related activities.

Instantiation. Based on the DPs, we instantiated our smart glasses-based service support system. All features were implemented on a native android application with the glass development kit based on android 4.4.2 (API 19). For the user interface we used Google Glass card designs that simplified the implementation. Because of the requirement to use voice recognition in the whole application, we further used custom layouts and handler. The interaction was mainly based on voice recognition. However, we implemented a fallback solution via the touch interface of Google Glass in case that voice recognition does not work (e.g. too noisy). *The derived design principles can be implemented on every smart glasses with a display and microphone (for voice-recognition).* The software was optimized to use the feature set of Google Glass. A more detailed discussion on hardware features was published by Niemöller et al. [10].

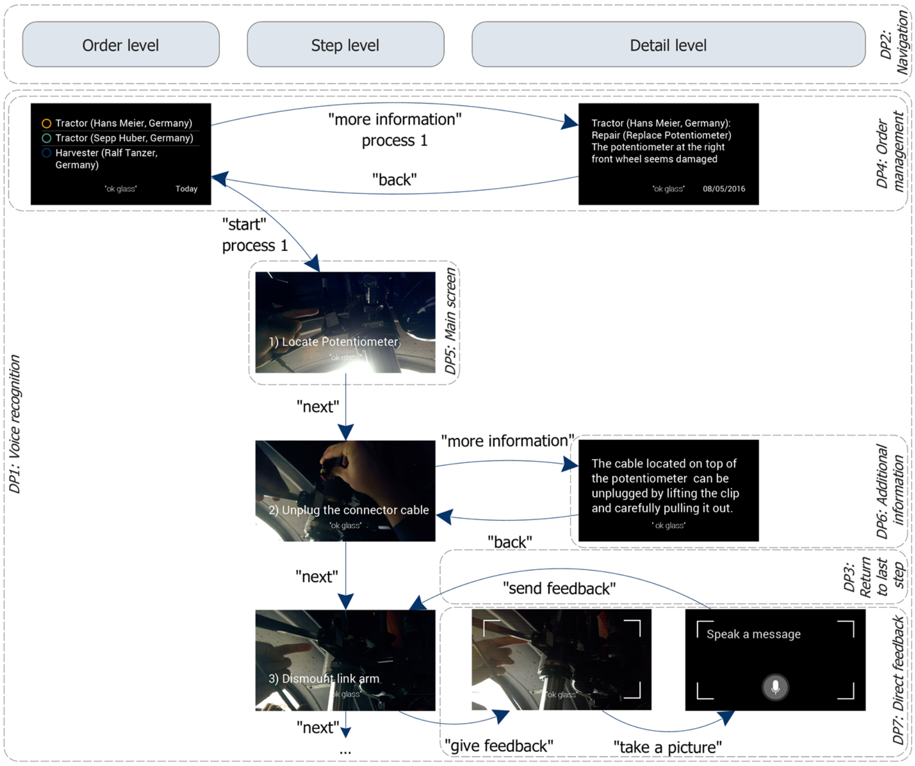


Figure 2. Screens of the smart glasses-based service support system (instantiating DP1-7)

The overall system is structured in three levels of navigation (cf. Figure 2). Following the DP of small menu navigation (DP2), we introduced an order level for the technician to look at the orders that has to be fulfilled. The second level is for the step-by-step guidance. The third one is for details and feedback. When technicians start using the service support system, they start at the order overview (fulfilling DP4). They get some general information about the orders they have to fulfill today (in our example three orders), what kind of machine it is, and the machine location. They get more information about the particular order when using the voice command *more information -> process 1*. In the order details the kind of order (e.g. maintenance, repair [3]) is specified as well as a short description on what the problem is and the timestamp when the order was commissioned (fulfilling DP5). With *back* they return to the order overview and start one order with *start -> process 1*. The step-by-step guidance turns up with the main screen. We designed the main screen as easy as possible with the number of the step, a short description on what to do and a picture in the background that illustrates what to do (fulfilling DP6). They get to the next step by the voice command *next*. When there is more information needed for a particular step, with the voice command *more information* (e.g. for further description of step in text form, video tutorials, pictures of tools) the system shows the attached information to that step (in our example it is a more detailed description; fulfilling DP7). With *back* the technician is taken back to the step. Later on, in step 4, the technician feels that there is something wrong. So, with *give feedback* the feedback module starts. First, a picture of the problem and, second, a message (via speech-to-text) can be recorded that is send to the backend (fulfilling DP8). With the voice command *send feedback* it is transferred and the system returns to the last step (fulfilling DP3).

The system architecture is implemented based on Google Glass with internet connection via Wi-Fi or Bluetooth (connected to a smartphone) and a backend-server that holds the data. For storing data locally (on Google Glass), we use a sqlite-database that is updated through a communication module talking to the backend-server. So, the administration of the processes and orders is done in the backend, whereas the communication module ensured that all data on Google Glass is up to date. We implemented the system in an agile approach and continuously discussed the results with the focus group members. Thereby, we evaluated the general feasibility of the system. Based on the final prototype, we evaluated the acceptance (cf. 5).

5 Evaluation

After having evaluated the single MRs and DPs using focus group meetings to inform our work (formative), we evaluated the system with a larger group of participants based on a demonstration of the prototype and survey (summative [32]). Goal of the evaluation was to proof the generality concerning different user groups and the applicability to real world problems (intention to use based on perceived usefulness and ease of use). We evaluated against the captured MRs. Addressing the five functional MRs, we asked the participants within our survey about whether they perceive the system to be useful (PU) for fulfilling their job. Addressing the non-functional MRs, we asked questions about the perceived ease-of-use (PEU). The wording of the four questions were taken from Venkatesh & Davis [35] and adapted to our scenario (Regarding PU: our system is considered as useful, if the technician feels empowered to fulfill her/his tasks in a higher quality and more efficiently [5, 6, 20] by being better informed [5, 23]). The two concepts are accompanied with one question about the behavioral intention to use (BI). This question gives an overall impression about whether they are willing to accept and use the system as demonstrated. Based on the TAM, both previous mentioned factors PEU and PU are meant to influence the BI. Figure 3 illustrates the statistical model. In total, 105 people participated in our survey. Most of them were male participants (86.7%) while all of them where between 16 and

60 years old (with an average of 31.4 years). Almost two out of three (62.9%) never had experience with smart glasses and the remaining third had experience only once or twice (35.2%). More than half of them (56.2%) are working in the agricultural machine and engineering industry. The remaining participants are working as agriculturalists (10.5%), in the IT industry (4.8%) or in other industries (22.9%).

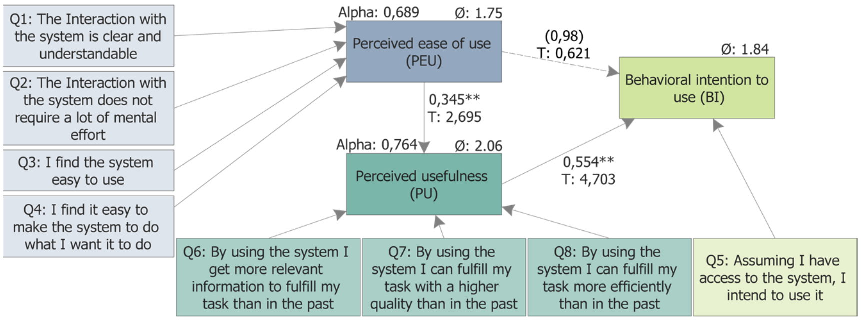


Figure 3. Summary of the evaluation based on TAM

In Figure 4, the factors perceived ease of use (left part) and perceived usefulness (middle part) are illustrated. With an average perceived ease of use of 1.75 which is in between *Highly agree* and *Agree* and no negative voting at all, a positive evaluation is given. So, the deduction that the non-functional MRs are evaluated positive is given. The perceived usefulness of the system with an average of 2.06 (Around *Agree*) is given as well. So, our functional MRs are evaluated positively as in average the participants perceive the usefulness of the system positive. The overall rating of the acceptance is evaluated through the factor behavioral intention to use. As the underlying TAM claim a correlation of behavioral intention to use and actual use, we argue that our system will be used in future if participants are evaluating their behavioral intention to use positive. Figure 4 (right part) also illustrates the results of the survey regarding the behavioral intention to use. With an average of 1.84 (Between *Highly agree* and *Agree*) and 77.1% of the participants rating positive on whether they intend to use the system if they have access to it, a positive feedback is given. Thus, we argue that people would accept and actually use the system in future. Indirectly, this evaluates our MRs and DPs positive as the system got positive feedback while being built on them.

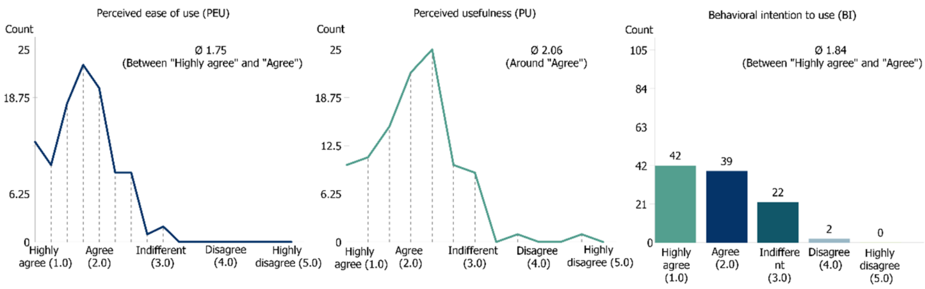


Figure 4. Evaluation results for PEU, PU and BI

Correlations During analysis, we conducted further calculations about correlation of the three factors perceived ease-of-use, perceived usefulness and behavioral intention to use as described in TAM [35]. We validated the significant positive correlation between the PEU and PU (Regression: 0.345, Significance: 0.008). Further, we found a significant positive correlation between PU and BI (Regression: 0.554, Significance: 0.000). However, the proposed positive correlation between PEU and BI could not be verified in our data (Significance: 0.536). One possible reason is the professional context the system should be used in. Thereby, the perceived ease of use alone does not necessarily lead to an adoption of a new system. Nevertheless, the indirect positive correlation through perceived usefulness still exists.

In sum, the evaluation of the system was positive. Thus, we were able to indirectly validate the DPs as the system was based on them and, thereby, the MRs. With our evaluation inspired by the TAM, we were also able to generate a forecast that with the positive feedback on the factor behavioral intention to use it is likely that people will actually use the system in future. The acceptance of the system was crucial for our evaluation strategy to validate the applicability to real world problems.

6 Discussion, Conclusion and Outlook

Conclusion. The effective use of emerging mobile IS can offer great opportunities to overcome current challenges in the domain of TCS. Due to the complexity of service systems engineering [18], guidance on how to design service support systems is necessary. To overcome this complexity and fill the research gap in design knowledge on smart glasses-based service support systems, we followed a DSR approach within this paper through, first, exploring the domain using triangulation and eliciting meta-requirements (RQ1), second, deriving design principles continuously working in an interdisciplinary team of practitioners and researchers (RQ2), and finally, evaluating the acceptance of our designed IS artefact (RQ3).

Novelty and Practical Relevance. We address a real-world problem that consists of the need for hands-free TCS service support through targeted information provision during work. At the same time, since smart glasses are still an emerging technology, little knowledge about the design of smart glasses-based service systems exist. During the evaluation phases, especially on the world's biggest agriculture technology fair, the demonstration of the prototype showed that the formulated design principles and their instantiation address the user's needs. The presented design principles can be transferred to other service domains as well as there are not TCS application specific, but specific to particular activities (e.g. specific design principle for step-by-step guidance-related activities.; hence, it is transferable to every domain where a step-by-step guidance is relevant). Another example are order-related activities which are transferable to every domain, where orders are relevant. Both from the point of practice and from theory, a transfer of the proposed design knowledge to other user groups offer new subjects of research, inter alia regarding value co-creation and new business models as e.g. the sale and delivery of smart glasses-based service support systems to the customer to enable self-services.

Theoretical Contribution. Regarding the theoretical contribution, this research work contributes to the methodological knowledge base of IS Design and Service Systems Engineering, and builds upon existing methods of DSR and findings in design of service systems [5, 6, 14, 23]. In DSR, a theoretical contribution is usually regarded to be in form of prescribing how a specific solution can be designed in order to solve a relevant real-world problem; often presented in form of design principles [37, 38].

Gregor and Hevner [16] argue that the instantiation itself contributes to the knowledge base as the demonstration of a novel artifact can be a research contribution that embodies design yet to be articulated, formalized, and fully understood. We position our work as a new solution, the hands-free information provision trough a smart glasses-based service support system, to solve an existing problem (need for service support due to complex and information-intense TCS processes). We explored the problem domain and formulated meta-requirements. They represent the conditions that should be met by a solution to provide the TCS with needed information while executing their processes without any need to interrupt their tasks.

Additionally, we contribute to the IS research knowledge base by instantiating the suggested methods by Sonnenberg and vom Brocke [33] while following the evaluation strategy proposed by Venable et al. [32] as enhancement of the classic DSR approach. Current discussions in service systems research argue for research to be embedded within a real-world scenario and call researchers to design novel service systems [18]. Hence, with our work, developed in a transdisciplinary team (IS research, service science, education and media psychology as well as practitioners from service providers, manufacturers and IT companies), we meet a research gap and the claim for evidence-based design research [18].

Limitations and Outlook. Although, we discussed our work with experts from two different sectors, the transfer of the design principles to other sectors have to be evaluated further because the TCS domain has a wide area of application. Based on the results of the evaluation cycles, we focused on non-functional requirements and the acceptance of the smart glasses system. Hence, (1) the transfer of additional functional requirements for handheld devices [5] have to be investigated further as their might occur some difficulties implementing for example invoicing functionalities due to the small screen (more natural on the tablet). (2) We have not conducted an evaluation regarding the actual economic and ergonomic benefit yet. Thus, the next step of our research is the evaluation of our instantiation in form of a field test [33] in the TCS of the agricultural technology company and the service provider for air-conditioning.

While analyzing the service processes and information needs of our focus group partner, we discovered that the knowledge base and access to information as basis for our service support system (e.g. handbooks, service manuals, trainings) differ in the considered companies based on different types of service providers and the complexity and variant diversity of their service objects. To sum up, our approach can be considered as first step with more research to come that is specifying new business models, value co-creation driven through new and enabling technology, the level of service integration and hybrid value creation regarding (a) it's influence on information needs and information access and (b) it's implications on service modularization and service design e.g. through integration of the customer.

References

1. Thomas, O., Loos, P., Nüttgens, M.: Hybride Wertschöpfung - Mobile Anwendungssysteme für effiziente Dienstleistungsproduktivität im technischen Kundendienst. Springer, Berlin, Heidelberg (2010)
2. Rossi, M., Tuunainen, K.V., Pesonen, M.: Mobile technology in field customer service. Business Process Management Journal. 13, 853–865 (2007)
3. Walter, P.: Technische Kundendienstleistungen: Einordnung, Charakterisierung und Klassifikation. In: Thomas, O., Loos, P., and Nüttgens, M. (eds.) Hybride Wertschöpfung. 24–41. Springer, Berlin, Heidelberg (2010)
4. Baines, T., Lightfoot, H., Smart, P., Fletcher, S.: Servitization of manufacture - Exploring the deployment and skills of people critical to the delivery of advanced services. Journal of Manufacturing Technology Management. 24, 637–646 (2013)
5. Matijacic, M., Fellmann, M., Özcan, D., Kammler, F., Nüttgens, M., Thomas, O.: Elicitation and Consolidation of Requirements for Mobile Technical Customer Services Support Systems. ICIS 2013. 1–16. Mailand, Italien (2013)
6. Ray, G., Muhanna, W.A., Barney, J.B.: Information Technology and the Performance of the Customer Service Process: A Resource-Based Analysis. MIS Quarterly. 29, 625–652 (2005)
7. Legner, C., Nolte, C., Nils, U.: Evaluating Mobile Business Applications in Service Maintenance Processes: Results of a Quantitative-Empirical Study. ECIS 2011. Paper 247 (2011)
8. Haritos, T., Macchiarella, N.D.: A Mobile Application of Augmented Reality for Aerospace Maintenance Training. 24th Digital Avionics Systems Conference. p. 5–B. IEEE (2005)
9. Henderson, S., Feiner, S.: Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. ISMAR 2009. 135–144. Washington, DC, USA (2009)
10. Niemöller, C., Metzger, D., Fellmann, M., Özcan, D., Thomas, O.: Shaping the Future of Mobile Service Support Systems – Ex-Ante Evaluation of Smart Glasses in Technical Customer Service Processes. Informatik 2016. 753-767, Klagenfurt (2016)
11. Rauschnabel, P.A., Ro, Y.K.: Augmented reality smart glasses: an investigation of technology acceptance drivers. Int. Journal of Technology Marketing. 11, 123–148 (2016)
12. Ernst, C.-P., Stock, B., dos Santos Ferreira, T.: The Usage of Augmented Reality Smartglasses: The Role of Perceived Substitutability (2016)
13. Hein, E.D.W., Rauschnabel, A.P.: Augmented Reality Smart Glasses and Knowledge Management: A Conceptual Framework for Enterprise Social Networks. In: Rossmann, A., Stei, G., and Besch, M. (eds.) Enterprise Social Networks: Erfolgsfaktoren für die Einführung und Nutzung - Grundlagen, Praxislösungen, Fallbeispiele. 83–109. Springer Fachmedien, Wiesbaden (2016)
14. Herterich, M.M., Peters, C., Uebernickel, F., Brenner, W., Neff, A.A.: Mobile Work Support for Field Service : A Literature Review and Directions for Future Research. WI2015. 134–148. Osnabrück (2015)
15. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Quarterly. 28, 75–105 (2004)
16. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. Management Information Systems Quarterly. 37, 337–355 (2013)
17. Walls, J.G., Widmeyer, G.R., El Sawy, O. a: Assessing Information System Design Theory in Perspective: How Useful was Our 1992 Initial Rendition. Journal of Information Technology Theory & Application. 6, 43–58 (2004)

18. Böhmann, T., Leimeister, J.M., Möslein, K.: Service Systems Engineering. Business & Information Systems Engineering. 6, 73–79 (2014)
19. Nüttgens, M., Thomas, O., Fellmann, M.: Dienstleistungsproduktivität - Mit mobilen Assistenzsystemen zum Unternehmenserfolg. Springer, Berlin (2014)
20. Agnihothri, S., Sivasubramaniam, N., Simmons, D.: Leveraging technology to improve field service. International Journal of Service Industry Management. 13, 47–68 (2002)
21. Däuble, G., Özcan, D., Niemöller, C., Fellmann, M., Nüttgens, M.: Design of User-Oriented Mobile Service Support Systems – Analyzing the Eligibility of a Use Case Catalog to Guide System Development. WI 2015. 149–163. Osnabrück (2015)
22. Becker, J., Beverungen, D., Knackstedt, R., Matzner, M., Müller, O.: Information Needs in Service Systems – A Framework for Integrating Service and Manufacturing Business Processes. HICSS 2011. 1–10 (2011)
23. Däuble, G., Özcan, D., Niemöller, C., Fellmann, M., Nüttgens, M., Thomas, O.: Information Needs of the Mobile Technical Customer Service - A Case Study in the Field of Machinery and Plant Engineering. HICSS 2015. 1018–1027. Manoa (2015)
24. Fellmann, M., Hucke, S., Breitschwerdt, R.R., Thomas, O., Blinn, N., Schlicker, M.: Supporting Technical Customer Services with Mobile Devices: Towards an Integrated Information System Architecture. AMCIS 2011. Paper 250. Atlanta (2011)
25. Metzger, D., Niemöller, C., Berkemeier, L., Brenning, L., Thomas, O.: Vom Techniker zum Modellierer - Konzeption und Entwicklung eines Smart Glasses. In: Thomas, O., Nüttgens, M., and Fellmann, M. (eds.), Smart Service Engineering. Springer Gabler. 193–213. Wiesbaden (2016)
26. Österle, H., Becker, J., Frank, U., Hess, T., Karagiannis, D., Krcmar, H., Loos, P., Mertens, P., Oberweis, A., Sinz, E.J.: Memorandum on design-oriented information systems research. European Journal of Information Systems. 20, 7–10 (2011)
27. Myers, M.: Qualitative Research in Business & Management. Sage, London (2009)
28. Dibia, V.: An Affective, Normative and Functional Approach to Designing User Experiences for Wearables. SSRN Electronic Journal. 1–12 (2015)
29. Gandy, M., Ross, D., Starner, T.E.: Universal design: Lessons for wearable computing. IEEE Pervasive Computing. 2, 19–23 (2003)
30. Smailagic, A., Siewiorek, D.: Application design for wearable and context-aware computers. IEEE Pervasive Computing. 1, 20–29 (2002)
31. March, S.T., Smith, G.F.: Design and natural science research on information technology. Decision Support Systems. 15, 251–266 (1995)
32. Venable, J., Pries-heje, J., Baskerville, R.: FEDS: a Framework for Evaluation in Design Science Research. European Journal of Information Systems. 1–13 (2014)
33. Sonnenberg, C., vom Brocke, J.: Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research. DESRIST 2012. 381–397 (2012)
34. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User acceptance of computer technology: a comparison of two theoretical models. Management Science. 35, 982–1003 (1989)
35. Venkatesh, V., Davis, F.D.: Theoretical extension of the technology acceptance model: four longitudinal field studies. Management science. 46, 186–204 (2000)
36. Trkman, P.: The critical success factors of business process management. International Journal of Information Management. 30, 125–134 (2010)
37. Kuechler, W., Vaishnavi, V.: A Framework for Theory Development in Design Science Research: Multiple Perspectives Science Research. Journal of the Association for Information Systems. 13, 395–423 (2012)
38. Sein, M.K., Henfridsson, O., Rossi, M., Lindgren, R.: Action Design Research. MIS Quarterly. 35, 37–56 (2011)

Towards Design Excellence for Context-Aware Services - The Case of Mobile Navigation Apps

Robin Knote[1], Matthias Söllner[1,2]

[1] University of Kassel, Information Systems, Kassel, Germany
{robin.knote,soellner}@uni-kassel.de
[2] University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland
matthias.soellner@unisg.ch

Abstract. To satisfy service customers and create unique value in a digitized world, companies must strive for exceeding customers' expectations of e-service experience by establishing high e-service quality. However, an increasing amount of e-services is performed by context-aware mobile technology, which is able to sense and react to changes in the user's environment. Although these context-aware services are able to address our personal needs and already determine our everyday live, knowledge on how to develop such services is sparse. In our study, we qualitatively compare three mobile navigation apps based on their user reviews in order to elicit first requirements and design approaches for e-service quality oriented design. Results show that well known e-service quality models are not fully applicable to the case of mobile navigation services.

Keywords: context-aware services, design requirements, e-service quality, user reviews

1 Introduction

In today's fast-paced digitized world, companies increasingly strive for creating unique value to customers in order to distinguish their own service offerings from those of competitors. Attracting, keeping and satisfying customers and making profit are becoming more difficult in times of low to zero marginal costs as one effect of product and service digitization. In the past, value was considered as ratio between service quality and cost [1]. However, Vargo and Lusch offer a new perspective on customer value by introducing the concept of value in use [2]. According to them, service users are both co-creators and judges of service value [1], which implies that how value is perceived is predominantly determined by customer's individual service experience [3]. In this regard, service quality, i.e., the discrepancy between expected and realized service experience [4], is the focal enabler of customer satisfaction [5, 6]. Whereas in a traditional human-to-human service encounter service employees are able to adapt interactively on customer reactions to improve service quality, an increasing amount of services are nowadays based on technology, which rely on electronic interfaces and predefined interaction procedures. In this context, service quality is usually conceptualized as e-service quality (ESQ), suggesting that interaction with e-services,

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Knote, R.; Söllner, M. (2017): Towards Design Excellence for Context-Aware Services - The Case of Mobile Navigation Apps, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 121-135

such as mobile applications or websites, is different to traditional human-to-human service encounters [7, 8].

With the rise of smartphones and the vision of ubiquitous computing, in which computers surround and adapt to the lives of humans without being obtrusive [9], the deterministic behavior of technology-based services, however, becomes increasingly turned upside down. Every modern mobile phone nowadays contains sensors to detect and adapt to changes of the user's physical or logical context, bringing context-aware mobile services a huge step towards non-determinism and ubiquity. Reflecting the new possibilities involved in the autonomous adaption to context information, such as time, location, speed, temperature, calendar entries, relevant news and (especially due to the success of smart watches) even health data like blood pressure and heart rate, adaptability of systems to users' real current needs is finally becoming possible. Context-awareness holds the potential to yield high-quality (i.e., perfect fit to individual users and their contexts) services that anticipate and hence satisfy customer needs in an entirely new manner. Against this backdrop, context-awareness can be considered a new and yet mainly unexplored paradigm that allows service and application designers to meet or even exceed customers' expected service experience in order to achieve high ESQ by design.

However, the question of how to design context-aware applications to meet requirements of high ESQ remains unanswered in scholarly and practical literature. In fact, even requirements that depict ESQ as design goal are sparse. This research gap may have manifold reasons. First, the definition and demarcation of service experience, service satisfaction, ESQ and related terms can be considered a conceptual chaos, since many schools of thought work towards many different directions. This leads to confusion in theory and practice when it comes to identifying the correct concept for a certain problem and its relation to adjacent ones. Second, ESQ as such has not be sensed as explicitly important for the design of (context-aware) services by practitioners, yet. Third, although many empirical data and a huge amount of measurement models exist on ESQ, no requirements or design principles for service and system development have been established based on this foundation. With this contribution, we aim to stress the importance of an ESQ-oriented system design approach. Our work is guided by the research question: *What are requirements and design approaches for high ESQ when developing mobile navigation services, as a subclass context-aware services?* We therefore identify, screen, cluster and analyze user reviews of three successful mobile navigation apps, as an important subclass of context-aware services, with regard to ESQ dimensions. In doing so, we aim at inductively identifying ESQ requirements and design approaches that help developers and requirements engineers to identify what makes the design of the mobile navigation services superior and what to consider when developing context-aware applications with ESQ as an overarching design goal.

The paper is structured as follows: In section 2, we give an overview on the peculiarities of context-aware services and introduce the concept of ESQ. Afterwards, our research approach is described in detail in section 3. Section 4 depicts the results of our study, which are then discussed in section 5. We conclude with a short summary, propose theoretical and practical contributions as well as limitations and next steps in section 6.

2 Background

2.1 Context-Aware Services

Following the vision of ubiquitous computing, one of the first ideas of applications that adapt to changes of their environment was introduced by Schilit and Theimer in 1993, who define the location, people and objects in proximity as well as their history as context of an application [10]. The question "Where you are, who you are with and what resources are nearby" thus characterize the context of an application [11, 12]. Ryan et al. define context as any collection of information about a system's environment, such as location, time, temperature and user identity [13]. Pascoe et al. state, however, that to define context is a complex task which cannot be done by simply listing possible context information [14]. Furthermore, Dey et al. conceptualize context as "any information about the user and the environment that can be used to enhance the user's experiences" [15]. In our work, we build up on abovementioned definitions and regard context as an attribute comprising all information about an entity's physical (e.g., location, temperature, humidity, etc.) and logical (e.g., calendar events, relevant news feeds, etc.) environment which can be captured by sensors. Furthermore, we focus on applications that are able to detect context (context detection) and react to this information by changing their behavior or functionality (context adaption). From a service science perspective, which focuses on the interaction and co-creation between entities within a service system to achieve a certain value, context-aware services play an increasingly important role [16, 17]. Glushko conceptualizes context-aware services as one field for future service design [18]. Nowadays, an increasing amount of technology-based services are aware of and adaptive to context, such as personal fitness, healthcare, disaster warning and navigation apps.

2.2 E-Service Quality

As a key driver leading to satisfied customers and economic growth, quality has its roots in marketing and consumer research, management science, engineering and operations. Golder et al. show that different understandings of quality have emerged in these academic fields [19]. The most pivotal perspective for our approach is obtained in marketing and consumer research which focuses customers' perception of quality. However, due to the heterogeneity of quality perspectives, "no universal, parsimonious, or all-encompassing definition or model of quality exists" [20]. One reason for the complexity of quality is that "objective quality may not exist because all quality is perceived by someone" [21]. In academic literature, the concept of perceived quality was proposed for services in manifold forms [8, 21–25]. Most existing approaches for analyzing and measuring service quality are based on the GAP model introduced by Parasuraman et al. [24] and SERVQUAL, a measurement instrument adopted from marketing [21]. Although they have massively shaped the research on service quality, these approaches have been established for services that encompass human actors only. However, due to digitization, many services are nowadays delivered and consumed through technology, which turns them into e-services. In the context of our study, we follow Riedl et al. who define e-service as “a business activity of value exchange that is accessible through an electronic interface” [26]. Parasuraman et al. correctly find, that service quality literature is dominated by people-delivered services whereas only a limited number of articles deal with customers’ assessment of ESQ, its antecedents and consequences [7]. However, there is a need for separately investigating ESQ, since customer evaluation of technology is different due to technology-based customer beliefs (e.g., technology-readiness) and technological peculiarities (e.g., server problems or connectivity issues) that may influence the quality perception process [7, 27]. Thus, Parasuraman et al. present a multiple-item scale for measuring ESQ called E-S-QUAL [7]. It is constructed and tested for measuring ESQ delivered by websites for online shopping. Despite the basic E-S-QUAL scale, which consists of 22 items in four dimensions (efficiency, fulfillment, system availability and privacy), they developed a scale for customers who have non-routine service encounters with the websites, which consists of 11 items in three dimensions (responsiveness, compensation and contact) called E-RecS-QUAL. Further models include zone-of-tolerance-based IS-SERVQUAL [28], SERVQUAL for E-commerce [29], eTailQ [30], WebQual [31, 32] and SaaS-Qual [33]. To consolidate all previous findings about ESQ and establish a common conceptual model, Blut et al. meta-analyze the most important empirical data [5]. In our work, we use their conceptual model as a starting point for developing coding categories. The meta-analysis conducted by Blut et al. reviews and summarizes a wide range of prior ESQ research and results in a conceptual model of four dimensions: website design, fulfilment, customer service and security. However, previous ESQ research also covered by Blut et al. has mainly focused on the interaction of customers with e-commerce websites [27]. For example, Parasuraman et al. define ESQ as “the extent to which a web site facilitates efficient and effective shopping, purchasing, and delivery” [7]. Thus, there is a huge research gap for investigating ESQ in a non-E-commerce context, such as context-adaptive services.

3 Research approach

In order to identify requirements and good design approaches for excellent mobile navigation apps, our approach is based on the method of Pegano and Maalej, who analyze user feedback from app stores and its impact on the user community [34]. For *data extraction*, we focus on the two biggest app stores: Apple’s AppStore and Google Play. We registered to an online service[1] for crawling app reviews from these platforms (by using e.g., Apple’s RSS feed generator) and converting them to comma-separated values (CSV) files for better analyzability. With this service, we extracted user reviews on top navigation apps on the US market: Google Maps, Waze and Navigon. According to download rate and overall user rating, these apps can be considered best-of-breed mobile navigation services. Our data sample covers the latest 1,500 user reviews (August 21, 2016; 500 reviews per app due to technical restrictions of the web service). The CSV data set contains the date, title and content of the review, nickname of

[1] https://heedzy.com/

reviewer, rating, reviewed app version and device of the reviewer. After extracting the reviews, *data analysis* was started by sorting out reviews with low to no informational value regarding requirements or design approaches for app development (e.g., ”Don’t download!!!” or “I love google maps”). Based on the ESQ dimensions and their definitions meta-analyzed by Blut et al. [5], we started building a category system for our endeavor, following the logic of qualitative content analysis [35]. However, since the meta-analysis mainly comprises studies that investigate ESQ with regard to e-commerce websites, not all dimensions conceptualized by Blut et al. are appropriate for our case of navigation services. This is mainly because (1) e-commerce providers use websites for information and facilitation of order processing whereas (context-aware) applications such as navigation apps are stand-alone e-services with backstage processes far behind the customer’s line of visibility, and (2) in e-commerce value is created through effective and efficient order processing and not by just using an electronic interface which, however, is the case for mobile applications. Hence, initial categorization was critically contrasted with other analyzed literature concerning ESQ. Categories must further fit to our research context and, following our understanding of context-adaptive services as mentioned above, reflect users’ perceptions of the app’s (a) context detection (i.e., gathering context data) or (b) context adaption (i.e., changing behavior based on context interpretation) abilities.

Thus, rearranging categories (i.e., ESQ attributes) from the initial framework was the first step: design was kept since we suggest it is important for app ESQ, especially when context changes (e.g., driving through a tunnel may cause the app to darken the screen to prevent dazzling). Following this logic, also information quality, ease of use, personalization and system availability have been kept from the design dimension, whereas e-commerce specifics (e.g., purchase process) were dropped. Furthermore, all categories regarding fulfilment and customer service were dropped from the initial framework, since they are by definition too e-commerce specific. Security and Privacy were kept since these appear crucial to context data detection. Reviewing prior research reveals that Benlian et al. is closer to our research context than other work in the field, since they provide not only ESQ items for user-to-provider but also for direct user-to-software encounters [33]. Some of these items are transferable to categories that are valuable for our approach. We hence added functional reliability and functional features (i.e., offline usability, search, etc.) as new categories. Search terms were defined, which we used to deductively classify the reviews into our categories. Since we conducted deductive content analysis, categories formed for coding qualitative data ought to be challenged by the actual data during analysis to evolutionarily emerge [35]. We hence inductively found one more category, network data usage. Final coding categories and definitions are presented in table 1. 583 helpful (i.e., related to one or more ESQ dimension and reflecting requirements and/or information about design approaches) user reviews were analyzed.

Table 1. Coding categories and definitions based on applicable ESQ dimensions

Source	Coding Category	Definition
Adapted from [5]	Design	Adaption of the user interface…
	Information quality	Actuality and accuracy of information…
	Ease of use	Intuitiveness of usability…
	Personalization	Fit to user's preferences…
	System availability	Stability during runtime…
	Privacy/Security	Perceived data protection…
Adapted from [33]	Functional reliability	Accessibility of desired functions…
	Functional features	Range of functionalities…
New	Network data usage	Data usage…

... when detecting or adapting to context.

4 Results

4.1 Design and information

In the following, we compare the results for the categories mentioned above by using quotes from user reviews as examples to highlight our interpretation. In this section, we focus on *Design, Information quality, Ease of use* and *Personalization.*

Design. In terms of general application design, Google Maps users clearly stress the superiority of how the app appeals. They found the app very well organized, accurate and convenient. Users tend to connote Google's design with quality, since it is clean and simple. For Navigon, however, the majority of the users accredit the app to have an "awkward" user interface design. Users do not perceive the user interface as contemporary and wish easier access to important menu points and settings to have better control over the service's functionality. *"[...] it seems that the company is relying on their exceptional map data to compensate for an exceptionally bad UI".* Waze, which is a navigation app based on crowdsourcing (i.e., users are able to mark traffic jams, red light cameras or police spots right in time), has been judged by users with both positive and negative evaluations almost equally. The comic style interface is rated as good and innovative by some reviewers, weird and confusing by others.

> *"Good design. Best navigation app ever!"*
>
> *"The UI is so bad that I find it easier to just quit the app and restart whenever I need to change something than to try to decipher the UI elements."*

Whereas Waze users do not explicitly judge design, Google users mainly appreciate the quality of the maps and the satellite imagers. However, when it comes to navigation (i.e., functionality for context adaption), some users find roads hard to identify since they are painted in shades of light and dark grey that are close to each other. Navigation design is considered easy and intuitive instead. Requirements are raised concerning unwanted rescaling of the map when using the 'find my location' button and icon sizes. For Navigon, also neither great concerns nor compliments are posted by the users. One minor requirement is to enlarge road signs (showing speed limit and other information on screen based on the user's location), making them easier to read. Also the Navigon app is perceived as superior in navigation mode. It is considered easy and reactive to users' driving behavior, as it displays the current speed, ETA and speed limit warnings.

> *"In general it works really well, gives accurate instructions and easy to follow graphical instructions that warn of upcoming changes. So I recommend it!!!"*

A minor requirement is raised regarding the text size of current speed, road name, ETA and speed limit in navigation mode. Waze users are also satisfied with the navigation design.

Information quality. Reviews on information quality are mostly related to the up-to-dateness of map material and the correctness of driving recommendations. User reviews on Google Maps reveal that location and road information as well as in-time traffic updates are very useful. Users perceive that the app adapts to their situation when suggesting alternative routes that are faster. Since map data is usually loaded via internet when needed and is updated regularly based on user feedback, traffic news and GPS patterns (e.g., for detecting traffic jam), map data is considered up-to-date.

> *"Real time traffic updates, route updates, and very clear, easy to follow travel instructions makes this app excellent."*

Although some users praise the actuality of its maps, the majority of Navigon users claim that map data is out of date. Furthermore, some users note that the navigation functionality is *"saddled with Microsoft's 'Here' map search which can never seem to find anything".* Waze users do not focus their reviews on road/map data up-to-dateness. However, some users posit that the quality and actuality of map and road data heavily relies on how active and efficient the crowd provides information. Thus, the crowd provides further information about the actual context and determines information quality.

> *"What Waze accomplishes is providing true real-time traffic and routing updates that are as good as the user community deserves. See, many use Waze for its superior alerts and routing capability over other GPS apps. However, it's really only as good as the user community is willing to make it"*

Ease of use. All three apps are described as very easy to use and intuitive by the majority of reviewers. As mentioned above, Navigon is perceived to have a complex menu organization which needs initial settling-in.

> *"Interface is straightforward although the depth of features sometimes requires digging around in settings and experimenting to get the app to do exactly what you want."*

However, all apps provide real-time adaption to context changes by displaying useful options. For example, Google Maps and Waze detect whether other drivers using the app are stuck in a traffic jam. If this is the case, it informs the driver and calculates alternative routes to bypass the overloaded road. Drivers can select a faster route by just clicking on the alternative route as it appears on screen.

Personalization. All apps provide the opportunity to store and recall personal waypoints for route calculation as well as individual points of interest. When the user is close to one of these points and selects the 'discover neighborhood' (or similar) option, the user is located and the distance to her personal point of interest is calculated. However, although it is possible, user reviews reveal that it is still difficult in Google Maps to set multi waypoint routes.

4.2 Functional features and reliability

As done above, in this section, we elaborate more on the findings for the categories *Functional features* and *Functional reliability*.

Functional features. Although the apps work on different mechanisms, all three services provide navigation and location functionalities. Google Maps and Navigon further provide the opportunity to use offline map data for navigation. However, both use different approaches for doing so: In Google Maps it is possible to zoom out of a map to a certain degree and save this section to the phone. Navigon works with downloadable region maps (e.g., for Europe, North America, etc.). In contrast, Waze needs a permanent internet connection in navigation mode to load map material. Although the majority of users appreciate the novel offline functionalities of Google Maps, users negatively rate the inconvenient way how offline maps are saved and used for navigation.

> *"It would really be nice, very helpful and much easier if I could save an entire state worth of offline map data instead of fidgeting with a clunky map screen I have to zoom and position so strategically to make sure I get all the roads, railroads and trails I need to have with me at all times."*

Google users especially enjoy that Maps offers up to three route alternatives with ETA which is calculated from length of the road, average personal speed, average speed on the road but also from traffic jams or construction zones detected. The user may choose from these alternatives, which they find simple and convenient. Further requirements are raised concerning a 'reverse option' to toggle start and destination and disturbing voice directions while using the phone for calls. Furthermore, Waze provides 'social' functionality, such as easily sending ETA to phone contacts while driving.

Functional reliability. Most Google Maps users describe the app extremely accurate with regard to the navigation functionality. However, complaints exist about lost GPS receptions in the middle of navigation and wrongly identified fastest routes.

> *"Quite often, it fails to find the fastest route to a destination. Consequently, I have taken to opening Waze and comparing routes. Around 50% of the time, Waze will find a faster way of getting where I need to go. What is keeping me from using that app full-time is their cartoon interface and ghastly color scheme. Google needs to vastly improve their algorithms for this app."*

Also the majority of Navigon users consider its navigation functionality accurate and reliable. Both Waze and Navigon successfully manage to route around traffic problems when internet access is available. User reviews further reveal that offline availability is becoming increasingly important for users.

4.3 Availability, Security and Network Data usage

This section is devoted to findings in the categories *System availability, Privacy/Security* and *Network data usage.*

System availability. As expected, most criticism on the apps is related to system availability. Although according to the users all three apps basically run stable, many users write reviews to complain about crashes and outages. Besides occasional instances in Google Maps and Waze, in which the apps randomly crash or lose network connection, more common concerns have been raised for the Navigon app. Users complain about data losses after updates.

"Navigon has screwed me again. The latest update wiped out all my Favorites. It also removes all the maps and requires individual downloads by state."

Privacy/Security. Navigon users do not raise concerns about data privacy and security. For both Waze and Google Maps, users complain about regularly established GPS connections even when location based services are turned off in phone settings and the apps are not running. A Waze user writes:

> *"Location Service Always Running even when waze is turned off [...] which is a lack of privacy. There is no need to have location service running all the time in background."*

However, users tend to hazard consequences of neglected privacy for the sake of functionality, how a Google Maps user describes:

> *"Yes, Google is probably collecting my usage data. At this point in my life I don't much care. The convenience of finding places and being accurately guided there true my hopes for anonymity."*

Network data usage. Waze users do not report on network data usage in online mode. Navigon users tend to rate mobile data usage positive.

> *"The data usage is very low - so low, it doesn't put a dent in my metered data (probably 200KB or less each 20 min trip)."*

Google Maps users, however, recently report about high data usage, which is seen as a severe issue.

> *"Recently I left this app on in the background while using walking directions and it used over 6 GB of data within hours, with no warning. This app needs a data warning, cut-off, or data usage metering like the Netflix app recently implemented."*

5 Discussion

As seen in the reviews, how users experience and evaluate context-aware e-services is highly dependent on individual situations and prior experiences. Many users compare service functionalities with what they know (i.e., formerly used apps) and thus build a state of expectation. This state is challenged when using a new app in order to form an evaluation of ESQ. Thus, it is necessary to consider ESQ as an important design goal for app and service development. Our results indicate that user reviews are a useful source of requirements and first hints for respective 'good practice' design approaches, which confirms conclusions of other scholars in the field of requirements engineering [34, 36]. We were able to extract 20 requirements from user reviews related to context-adaption or context-detection, which thus apply to our understanding of context-aware services. Furthermore, 8 of these requirements can be linked to design approaches that could also be identified in our work. Table 2 provides an overview on the requirements (R) and – if existent – design approaches (D) derived from the user reviews and mapped to context-aware services.

Results regarding the *design* reveal that users of navigation apps prefer a simple and straightforward design which is easy to read and handle. Convenience is especially desired since context-aware applications are mostly used to obtain additional information while concentrating on a real world activity, such as driving. The quality of map data in terms of picture resolution and up-to-dateness is crucial for apps which provide location or navigation functionalities. Basic navigation should be enriched with a choice set of alternative routes. Other requirements in this category relate to map scaling and readability. *Information quality* entails the necessity to continuously update context information by balancing requests and informational value. Further, users value accurate context information which can be established by relying on multiple sources of context data. In case the context-aware application relies on crowdsourcing mechanisms (e.g., Waze for navigation), participation must be enabled and should be rewarded to raise information quality. To increase the *ease of use*, important menu points and settings should be conveniently accessible for the reader (e.g., while driving). In case location search and navigation is used in a context-aware application, users should be able to set personal waypoints and find points of interest relevant to their current needs in order to increase *personalization.* Furthermore, the use of location based services also allows for the implementation of *functional features* which users consider valuable. These are offline use of data (e.g., downloading a map for offline navigation), multi waypoint routes, reverse option for start and destination, route alternatives and precedence to more important applications. Especially the last feature is of interest for future work on context-aware services because it necessitates an integrated view on different services, their interaction and prioritization. Users further require *system availability* to be stable twenty-four-seven. Also the application itself should run without errors or crashes in any context. In order to establish *Privacy and security,* users desire control over when the app gathers context data. Although US users tend to neglect privacy and security concerns to obtain outstanding functionality, other (e.g., cultural) contexts may have formed different understandings, which have to be considered, since individual, organizational and cultural factors are suggested to moderate the effects on ESQ [5]. In addition, *network data usage* should constantly rely on few mobile data traffic and prompt a warning in case data capacity is close to user's data plan limit.

Table 2. Requirements and design approaches derived from user reviews

Requirement/Design approach
Design
R1: Simple Design: Context-adaptive applications should have a clean and simple design. Texts should be easy to read.
D1: Use modern and straight design with color contrasts. Refrain from using comic style.
R2: High quality map data: In case the app provides location detection and navigation functionality, high quality map data (i.e., resolution and up-to-dateness) should be provided. Zooming shout be possible on maps.
D2: Refrain from using similar colors for different elements (e.g., light grey for street, dark grey for houses). Provide high-resolution and up-to-date (also continually updated) map material.
R3: Prevent map from unwanted rescaling: If location detection on maps are used, users should have control over the scaling. Unwanted rescaling should be prevented.
R4: Easy to read additional context information: In case context changes, users should receive an easy to read but non-disturbing information (e.g., speed limit warning).
Information quality
R5: Continuous context updates: Users should at any time obtain information about the context which is up-to-date.
D5: Balance between information distribution and server requests. Load new context data (e.g., map data) only when needed.
R6: Accurate context information: Users should receive accurate context information at any time.
D6: Use multiple sources of context data, such as built-in sensors, user feedback, news and information from a crowd of users. Use effective algorithms for calculating fastest routes.
R7: Crowd participation: In case of sharing context information with the crowd, users should be able to easily participate.
D7: Give users easy access to crowd activities. Motivate crowd to participate.
Ease of use
R8: Easy access to important menu points: Important menu points and settings should be easy to access by the user.
D8: Provide short menu on top screen with most important settings which is easy to access while concentrating on real-world activities (e.g., driving)
Personalization
R9: Personal waypoints: In case maps are used, the user should be able to store and recall personal waypoints such as points of interest.

R10: Points of interest nearby: In case location search and/or navigation is used, users should receive recommendations to points of interest nearby based on their interests and current needs. Users should receive information and may let the app navigate to the respective destination.
Functional features *R11: Multi waypoint routes:* In case navigation is used, users should be able to set routes with more than one waypoint. *R12: Reverse option for start and destination:* If route planning or navigation is provided, users should be able to easily toggle start and destination. *R13: Route alternatives:* If navigation is used, users should choose from different route alternatives. ETA should be provided for the routes, too. Route alternative selection should also be available just-in-time, especially when context changes (e.g., traffic jam dissolves). *D13:* Calculate different route alternatives and ETA and provide the top two to three to the user. Users can then select the most appropriate route for them by clicking on one alternative. *R14: Precedence to more important applications:* In case apps are started that need access to hardware resources (e.g., incoming phone calls need access to speaker and microphone), users should be able to configure the app so it gives precedence to these applications (e.g., by stopping voice directions for navigation). *R15: Offline data:* Users should be able to download static data (such as map material or points of interest) to use them with context functionality (e.g. navigation) when no network access is available.
System availability *R16: 24/7 service availability:* Users should be able to access the service whenever they want. *D16:* Optimize network, system and service stability. Come as closest to 100% availability as possible. *R17: App stability:* Users should be able to run all functionalities of the app without failures, crashes and data losses in any context. Updates of the app should not lead to loss of data.
Privacy/Security *R18: Limit context detection:* The app should not gather context information from users, when it is not directly needed for the functionality the user executes. Users should further be able to completely turn off context detection.
Network data usage *R19: Use few network data:* In case no WiFi is available, the app should not use more mobile data capacity than needed for the current functionality. *R20: Data exhaustion warning:* Users should be warned in case network data capacity is high and close to the limit of the user's data plan.

Table 2. (cont.)

6 Conclusion and Outlook

In this work, we investigated how ESQ as main determinant of e-service satisfaction is established for mobile navigation services. By conducting a qualitative content analysis of user reviews for mobile navigation apps, we identified requirements and first design approaches to meet or even exceed customer's expected service experience and thus achieve high ESQ. However, future studies, qualitative as well as quantitative, must show if our categories comprehensively explain the effects of ESQ for context-aware (navigation) services and if requirements and design approaches identified in this study are applicable in a practical development context. Positioning ourselves in the field of ESQ, we are, to the best of our knowledge, the first who not only investigate ESQ from a purely empirical but also from a design perspective. Since reviewing and eliciting requirements from user comments in app stores is an upcoming approach in the field of requirements engineering [34, 36], building up on previous methodical work seems appropriate for our endeavor. However, since our sample is limited we can only shed light on a subset of possible ESQ requirements and design approaches that possibly hide in the mass of user reviews. Future research should investigate other app stores, such as the windows store, as well as cultural contexts other than the US to triangulate our findings. From a theoretical viewpoint, we propose to challenge ESQ theory-building approaches by introducing an inductive approach of qualitative category development. We suggest that ESQ as highly individual, perception-based construct cannot be fully explained by quantitative research only. We show that even a consolidated model of ESQ [5] cannot fully be applied on each case of interest. By further investigating more cases of context-aware services, (e.g., disaster warning systems), a theory covering ESQ for context-aware services will emerge. Furthermore, the accumulation of empirically grounded requirements and design approaches for ESQ-oriented design of such services can be considered as nascent theory of design and action [37]. We further contribute to requirements, software and service engineering practice, since we elaborate on what to consider when designing mobile navigation services with regard to ESQ. User reviews increasingly become an important source for evaluating design approaches and identifying opportunities for improvement. By comparing three good practice designs and consolidating both positive and negative design aspects, we elicit requirements that address contemporary customer needs and enable software and service engineers to establish high ESQ by design.

References

1. Sandström, S., Edvardsson, B., Kristensson, P., Magnusson, P.R.: Value in use through service experience. Managing Service Quality: An International Journal 18, 112–126 (2008)
2. Vargo, S.L., Lusch, R.F.: Evolving to a New Dominant Logic for Marketing. Journal of Marketing 68, 1–17 (2004)
3. Bitner, M.J.: Servicescapes. The Impact of physical surroundings on customers and employees. Journal of Marketing 56, 57–71 (1992)

4. Bitner, M.J., Zeithaml, V.A., Gremler, D.D.: Technology’s Impact on the Gaps Model of Service Quality. In: Maglio, P.P., Kieliszewski, C.A., Spohrer, J.C. (eds.) Handbook of Service Science, pp. 197–218. Springer US, Boston (2010)
5. Blut, M., Chowdhry, N., Mittal, V., Brock, C.: E-Service Quality. A Meta-Analytic Review. Journal of Retailing 91, 679–700 (2015)
6. Xu, J., Benbasat, I., Cenfetelli, R.T.: Integrating Service Quality with System and Information Quality. An Empirical Test in the E-Service Context. MIS Quarterly 37, 777–794 (2013)
7. Parasuraman, A., Zeithaml, V.A., Malhotra, A.: E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality. Journal of Service Research 7, 213–233 (2005)
8. Zeithaml, V.A., Parasuraman, A., Malhotra, A.: Service Quality Delivery through Web Sites. A Critical Review of Extant Knowledge. Journal of the Academy of Marketing Science 30, 362–375 (2002)
9. Landay, J.A., Borriello, G.: Design patterns for ubiquitous computing. IEEE Computer 36, 93–95 (2003)
10. Schilit, B.N., THeimer, M.M., Welch, B.B.: Customizing Mobile Applications. Proceedings of USENIX Symposium on Mobile & Locaiton-Indendent Computing, 129–138 (1993)
11. Sitou, W.O.: Requirements Engineering kontextsensitiver Anwendungen. München (2009)
12. Schilit, B.N., Adams, N., Want, R.: Context-Aware Computing Applications. IEEE Workshop on Mobile Computing Systems and Applications (1994)
13. Ryan, N., Pascoe, J., Morse, D.: Reality Fieldwork: the Context Aware Archaeological Assistant. In: Gaffney, V., van Leusen, M., Exxon, S. (eds.) Computer Applications in Archaeology (1997)
14. Pascoe, J., Ryan, N., Morse, D.: Issues in Developing Context-Aware Computing. Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing, 208–221 (1999)
15. Dey, A., Abowd, G., und Wood, A.: CyberDesk: a framework for providing self-integrating context-aware services. Knowledge-Based Systems 11, 3–13 (1998)
16. Böhmann, T., Leimeister, J.M., Möslein, K.: Service Systems Engineering. A Field for Future Information Systems Research. Business & Information Systems Engineering 6, 73–79 (2014)
17. Vargo, S.L., Akaka, M.A.: Service-Dominant Logic as a Foundation for Service Science. Clarifications. Service Science 1, 32–41 (2009)
18. Glushko, R.J.: Seven Contexts for Service System Design. In: Maglio, P.P., Kieliszewski, C.A., Spohrer, J.C. (eds.) Handbook of Service Science, pp. 219–249. Springer US, Boston (2010)
19. Golder, P.N., Mitra, D., Moorman, C.: What Is Quality? An Integrative Framework of Processes and States. Journal of Marketing 76, 1–23 (2012)
20. Reeves, C.A., Bednar, D.A.: Defining Quality: Alternatives and Implications. Academy of Management Review 19, 419–445 (1994)

21. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: SERVQUAL: A Multiple-Item Scale for Measuring Customer Perceptions of Service Quality. Journal of Retailing 64 (1988)
22. Grönroos, C.: Strategic Management and Marketing in the Service Sector. Research Report Number 8, Swedish School of Economics and Administration (1982)
23. Grönroos, C.: Service Quality. The Six Criteria of Good Perceived Service Quality. Review of Business 9, 10–13 (1988)
24. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: A Coneptual Model of Service Quality and its Implilcations for Future Research. Journal of Marketing 49, 41–50 (1985)
25. Cenfetelli, R.T., Benbasat, I., Al-Natour, S.: Addressing the What and How of Online Services: Comparing Service Content and Service Quality for E-Business Success. Information Systems Research 19, 161–181 (2008)
26. Riedl, C., Leimeister, J.M., Krcmar, H.: Why e-Service Development is Different. A Literature Review. e-Service Journal 8, 2–22 (2011)
27. Collier, J.E., Bienstock, C.C.: Measuring Service Quality in E-Retailing. Journal of Service Research 8, 260–275 (2006)
28. Kettinger, W.J., Lee, C.C.: Zones of tolerance: Alternative scales for measuringinformation systems service quality. MIS Quarterly 29, 607–623 (2005)
29. Gefen, D.: Customer Loyalty in E-Commerce. Journal of the Association for Information Systems 3, 27–51 (2002)
30. Wolfinbarger, M., Gilly, M.C.: eTailQ: Dimensionalizing, measuring, and predicting etail quality. Journal of Retailing 79, 183–198 (2003)
31. Loiacono, E.T., Watson, R.T., Goodhue, D.L.: WebQual: An instrument for consumer evaluation of Web sites. International Journal of Electronic Commerce 11, 51–87 (2007)
32. Barnes, S.J., Vidgen, R.: An evaluation of cyber-bookshops: The WebQual method. International Journal of Electronic Commerce 6, 11–30 (2001)
33. Benlian, A., Koufaris, M., Hess, T.: Service Quality in Software-as-a-Service. Developing the SaaS-Qual Measure and Examining Its Role in Usage Continuance. Journal of Management Information Systems 28, 85–126 (2012)
34. D. Pagano, W. Maalej: User feedback in the appstore: An empirical study. In: 21st IEEE International Requirements Engineering Conference (RE), pp. 125–134 (2013)
35. Krippendorff, K.H.: Content Analysis. An Introduction to Its Methodology. SAGE, Los Angeles (2013)
36. Maalej, W., Nayebi, M., Johann, T., Ruhe, G.: Towards data-driven requirements engineering. IEEE Software 33, 48–54 (2016)
37. Gregor, S.: The Nature of Theory in Information Systems. MIS Quarterly 30, 611–642 (2006)

Recombinant Service System Engineering

Daniel Beverungen, Hedda Lüttenberg, Verena Wolf

Paderborn University, Department Wirtschaftsinformatik, Paderborn, Germany
{daniel.beverungen,hedda.luettenberg,verena.wolf}@upb.de

Abstract. Although many methods have been proposed for engineering services and customer solutions, most of these approaches give little consideration to recombinant service innovation. In an age of smart products and smart data, we can, however, expect that many of future service innovations need to be based on adding, transferring, dissociating, and associating existing value propositions. The purpose of this paper is to outline what properties constitute recombinant service innovation and to identify if current service engineering approaches fulfill these properties. Based on a conceptual in-depth analysis of 24 service engineering methods, we identify that most methods focus on designing value propositions instead of service systems, view service independent of physical goods, are linear or iterative, and incompletely address the mechanisms of recombinant innovation. We discuss how these deficiencies can be remedied and propose a first conceptual model of a revised service system engineering approach.

Keywords: Service engineering, recombinant innovation, (product-)service system, literature analysis, new service development

1 Introduction

The structured design of value propositions—also referred to as Service Engineering or as Product-Service Systems (PSS) Engineering [1–3]—has been a focal area of the Service Science discipline since the 1980s. Ever since, a plethora of methods has been proposed for designing 'services' or 'customer solutions' that consist of services, products, and information technology [1]. Against the properties of 'service' as the basic unit of exchange [4], we will refer to all these methods as 'service engineering' here for short. Most service engineering methods prescribe service design as a top-down engineering process that spans from idea management to introduction of a value proposition onto the market. Subsequently, service is co-created by service providers and service customers, thereby generating value-in-use for the stakeholders involved.

While the relevance of service engineering has increased [5], our understanding of service engineering has also shifted conceptually. In particular, the advent of smart products has enabled companies to offer value propositions that rely on context-specific field data that are made available in real-time. Discussed under the headword 'Internet of Things' or 'Internet of Services', these trends usher a new era of (smart) service systems engineering that is increasingly focused on designing integrated conglomerates

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Beverungen, D.; Lüttenberg, H.; Wolf, V. (2017): Recombinant Service System Engineering,
in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung
Wirtschaftsinformatik (WI 2017), St. Gallen, S. 136-150

of products, services, and information technology, which jointly provide value propositions based on which service and value-in-use are co-created [4], [6], [7].

However, as opposed to a considerable body of knowledge on service engineering, the value and applicability of the available methods for the era of smart service is questionable for two reasons. First, many of the available methods seem complex, over-engineered and overwhelmingly cumbersome, and require large investments to be made before a value proposition can be offered on the market [3], [8]. Second, most approaches implicitly assume an inside-out perspective that is based on defining (modular) value propositions that service providers offer to their clients [8], [9]. In contrast, the progressing availability of smart products and smart data suggests that many future innovations will be recombinant instead [10]. A recombinant innovation is not designed and brought to market by means of a top-down engineering process, but is developed by combining existing resources and solutions supplied by different stakeholders, and filling the gaps between them to co-create innovative value-in-use in a service system.

The purpose of this paper is to conceptualize recombination as a type of service innovation and—based on this conceptualization—to assess the suitability of existing service engineering methods to foster this type of innovation. We answer the following research question: To what extend do current service engineering methods support recombinant service innovation?

The remainder of the paper is structured as follows. In Section 2, we review and discuss related literature on service engineering, new service development, and (product-)service system engineering, as well as literature on service innovation and service modularization. In Section 3, we explain and justify our research method that includes a literature review and a conceptual analysis of service engineering methods. In Section 4, we report the findings of our conceptual analysis. In Section 5, we propose design principles and present a conceptual service system engineering approach that implements these design principles. Section 6 concludes the paper.

2 Related Research on Service Engineering and Innovation

2.1 Developing/Engineering (Product-)Service (Systems)

The first approaches covering the development of services were published under the banner of "New Service Development" (NSD) in the Anglo-American literature of the 1980s [11]. Johnson et al. [12] outline why "NSD research mirrors that in NPD" (New Product Development) and focuses on success factors, which "address *what* should be done, not *how* it should be done" [12] (emphases contained in the text). NSD mainly focuses on particular issues in service development, e.g. quality [13], [14], prerequisites for services [13], service blueprinting [15], or enablers for service development [12]. The approaches often contain frameworks or (partial) processes without presenting detailed methods or tools for service development [12]. Also, they often focus on a service management or service marketing perspective [11], [13], [16].

In parallel to NSD, another research stream started in the 1990s, transferring know-how from engineering disciplines and software development to service development

[5]. Standardized process models, methods, and tools for product and software development were analyzed and adapted for service development [5]. The aim was to build on advantages of engineering processes like improved efficiency, reduced development time and costs, and increased quality for service development [11].

A center of activities in this research stream was in Germany, where the term "service engineering" was used since the mid-1990s [5]. Here, several initiatives, conferences, and publicly funded projects were initiated since 1994 to strengthen the research activities and competences in structured service development [5]. From the funding program *Dienstleistungen für das 21. Jahrhundert* (Services for the 21st Century), service engineering emerged as an independent focus topic [5].

Several process models for service engineering have been designed in papers and several PhD theses [17–19]. Early approaches feature three to seven steps that can be repeated iteratively. These approaches have close references to product engineering approaches and, therefore, consider service as a product without taking into account other aspects, such as organizational or social impacts [2], [11], [19], [20].
More recent research extends the point of view from designing a value proposition to designing a service system. Scheuing and Johnson [16] already highlight the necessity to convert "the new service concept into an operational entity". Klein [17] develops a systems engineering approach based on considering the service engineering system as a social system. Becker et al. [3] identify different conceptualizations of product-service systems. Böhmann et al. [2] "conceptualize a service system as a socio-technical system that enables value co-creation guided by a value proposition"; it includes "not only data and physical components, but also layers of knowledge, communication channels and networked actors" [2]. Service is a "collaborative process creating context-specific value" and can be supported by information systems [2]. Engineering service systems comprise defining service architectures (i.e., modules of a service system and their interactions), designing interactions in service systems, and mobilizing human, physical, and information resources [2].

2.2 Recombinant (Service) Innovation

Innovation in general can be defined as a discontinuous change and describes a new solution or renewal of an existing solution [21]. As opposed to mere invention, innovation has practical or commercial value [22].

The extant literature conceptualizes six innovation processes [23] that can be either planned, intentional, or unintentional, which emphasizes an innovation's emergent character [24]. These innovation processes are: radical innovation, improvement innovation, incremental innovation, ad hoc innovation, recombinant innovation, and formalization innovation. In theory, most innovations are based on some sort of recombination [22], since hardly any innovation cannot be deduced from prior known building blocks [25]. Therefore, we will focus on recombinant innovation here.

Recombinant innovation relies on combining existing elements to design new services or to generate a new relationship between previously uncombined components [26]. It has been claimed to be a role model for service innovation [24] that can lead to incremental improvements as well as radical changes [27] in service systems.

The four basic operations of recombinant innovation are summarized and visualized in Figure 1. They can be concatenated to build more complex innovation patterns.

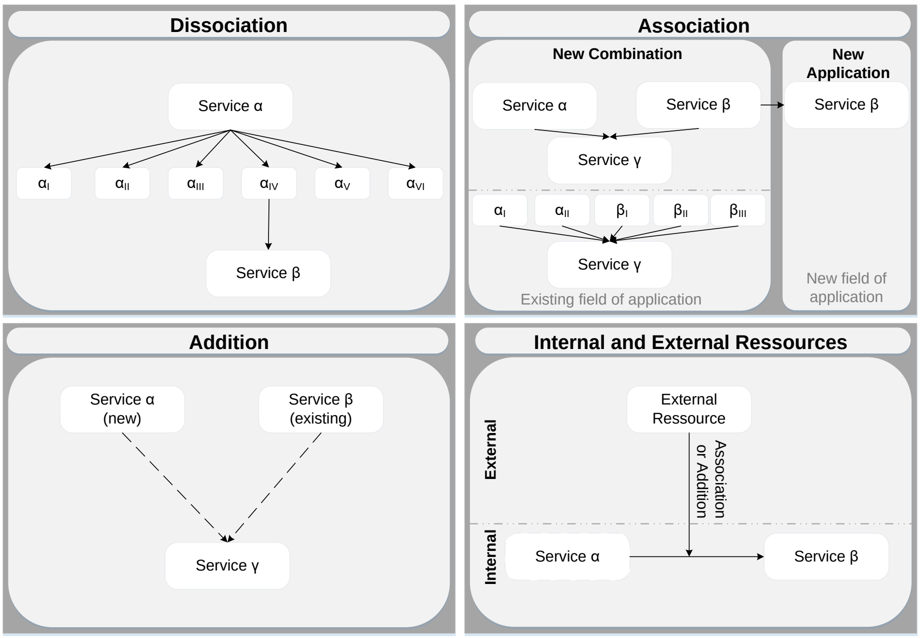


Figure 1. The four basic operations of recombinant innovation

Dissociation and association are two basic principles of recombinant innovation [28]. Association refers to designing a new value proposition by combining (or "associating") two or more existing services. Theoretically, any component can be recombined with any other component [26]. This indicates that the number of new combinations is a combinatorial function of the number of existing ideas [29]. An existing service can also be transferred to another context for which it was initially not designed [21]. Dissociation refers to designing a new value proposition by splitting up an existing one, isolating certain characteristics or a subset of operations, categorizing them, and turning certain elements into marketable services [30]. Services that have been split up into elements can be combined or integrated with other elements that were unconnected before [27]. Another principle of recombination is the addition of new value propositions [29].

Knowledge that is recombined can be drawn from internal and external resources. Internal resources refer to the capability to recombine a company's internal procedures in storing, retrieving, and processing knowledge [23]. Externally, firms retrieve knowledge through their relation with customers, suppliers, and other stakeholders that are involved in a service system [23]. Their relations give them access to valuable resources that cannot be generated internally. If resources possessed by the involved parties are similar, knowledge can be transferred efficiently leading to innovations, which however are rather incremental [31]. Integrating distant resources can lead to innovative breakthroughs, but presuppose that the actors can overcome cognitive disparities to absorb new knowledge efficiently.

Recombinant innovation relies on assumptions. First, it is assumed that a service can be broken down into clearly identified and defined elements [31]. Second, it is assumed that firms have the ability to maintain variety [27]. Third, recombinant innovation operates through continuous and cumulative creation of knowledge [30]. Managing this knowledge is complex [30], since successful recombinant innovation requires formalization and is therefore often built on a modular architecture [32]. This architecture allows the systematic reutilization of elements, which also leads to major resource savings [30]. Fourth, recombinant innovation requires certain competences of the agents, development work, and creativity [23].

2.3 Mass Customization and Modularization of Value Propositions

Mass customization [33] is a well-known strategy to efficiently deal with heterogeneous customer demand, based on configuring (seemingly) individualized value propositions that are composed of pre-defined modules.

Most of these methods include a service engineering process in which an initial set of goods, services, and IT modules are designed, often by defining a catalogue of items that are for sale [34], [35]. The design of modules is based on principles of strong cohesion and loose coupling. A crucial part of the engineering process is to specify modules and configuration rules with a (semi-)formal modeling language [35], [36]. The service engineering process is concluded with publishing a modular service architecture [37] that specifies the available components independent of specific customer requests.

In order to develop a value proposition for a particular customer, a service provider has to identify the needs, wants, and demands [38] of a particular customer, and configure a value proposition accordingly. The configuration process is based on fitting a subset of pre-defined modules together, allowing service providers to offer (seemingly) individual value propositions that fit a particular customer's demand.

3 Research Method

We performed a literature review to elicit what properties constitute recombinant innovation and analyzed which service engineering methods fulfill these properties.

The literature review was performed in line with the guidelines proposed by Webster and Watson [39]. After completing an informal screening phase, we compiled service engineering methods in several electronic libraries by applying German and English search strings. The literature research was conducted in the online data bases "Business Source Complete (via EBSCO Host)", "Association of Information Systems Electronic Library" (AISEL), and "Scopus". In addition, an extensive search was conducted in the journal Service Science, a local university library, and an inter-library loan system. Additional papers were compiled in a forward and backwards search to pinpoint articles that remained unidentified in the initial phase [39]. The search identified 24 methods.

Based on the literature we developed a concept matrix with eight dimensions. The dimensions were derived from the key properties of recombinant innovation and service engineering. Three coders individually used the concept matrix to analyze the identified service engineering methods. The initial inter-coder reliability [40] was measured using average pairwise percent agreement (A_0= 0.861), Fleiss' Kappa (κ=0.676), average pairwise Cohen's Kappa (κ=0.676), and Krippendorff's Alpha (α= 0.677). As all values are above the critical value of α_{min}= 0.667, concordance between the agents regarding the identified criteria can be assumed. Subsequently, we conducted a workshop to discuss and remedy conflicting assessments, until saturation was reached.
From the resulting concept matrix, a conceptual analysis was performed to identify the current state and research perspectives for service system engineering. Conceptual research can be used to initiate theory development and to assess and enhance theory [41]. Based on the analysis, we identify design principles that need to be considered by service system engineering methods and develop a conceptual approach as a prototype that communicates the design principles.

4 Conceptual Analysis of Service Engineering Methods

We use a concept matrix to provide a systematic review of service engineering methods [39]. While the matrix identifies the completeness of each method vis-á-vis theoretical concepts, any gaps and other topics are identified on the population level [39].

Our concept matrix is built on four types of constructs (Table 1). First, we identify addition, dissociation, and association as basic operations of recombinant innovation. Addition refers to recombining an existing value proposition with a novel characteristic. While each service engineering approach adds a new value proposition, we excluded approaches that did not explicitly identify preexisting value-propositions. Dissociation refers to a value proposition that is disaggregated into sub-components that are marketable themselves and/or can be combined with other modules adjacently. Association is a new combination of pre-existing value propositions or transfer of value proposition in a context for which this value proposition was not explicitly designed. Importantly, we submit that assembling pre-defined modules into (seemingly) unique solutions—as often done in modular service architectures—is not association in terms of a recombinant innovation, since modular service architectures often assume a finite solution space. Instead, we consider association to happen at design time, establishing new composite modules and/or new configuration rules.

Second, we identify if a service engineering method applies to designing a value proposition or if it is focused on designing a service system as socio-technical system for value co-creation. Since recombinant innovation often represents a combination of internal and external resources [42], service system engineering should identify the resources contributed by stakeholders in a service system early on.

Third, the type of process is identified as linear, iterative, or prototyping [43]. Linear models are characterized by discrete and consecutive process steps [43]. Iterative models exhibit multiple repetitions of the involved activities [43]. In prototyping models, a value proposition is refined by means of prototypes that communicate design ideas and explore the solution space, as proposed in Design Thinking [44].

Fourth, we identify if a service engineering method comprises the design of value propositions solely or if it also covers the design of value propositions in combination with physical goods. This combination of value propositions with physical goods results in Product-Service System (PSS), which allow the creation of new business models and added values for the customers [45].

Table 1. Conceptual analysis of service engineering methods, in chronological order

	addition	dissociation	association: new combination	association: transfer	internal / external resources	model scope	model type	type of solution
Scheuing et al. [16]	-	-	-	-	x	Value Prop.	linear	Service
Shostak et al. [15]	-	-	-	-	-	Value Prop.	iterative	Service
Edvardsson et al. [13]	-	-	-	-	-	System	linear	Service
Ramaswamy [14]	-	-	-	-	-	Value Prop.	linear	Service
Schwarz [46]	-	-	-	-	-	Value Prop.	linear	Service
DIN Fachbericht 75 [20]	x	x	x	-	-	Value Prop.	linear	Service
Jaschinski [19]	x	x	x	-	x	System	iterative	Service
Johnson et al. [12]	-	-	-	-	-	System	iterative	Service
Schreiner et al. [47]	-	-	-	-	-	Value Prop.	linear	Service
Meiren et al. [11]	-	-	-	-	-	System	linear	Service
Morelli [48]	-	x	x	x	-	System	iterative	PSS
Schneider et al. [43]	x	-	-	x	x	Value Prop.	linear	Service
Kunau et al. [49]	x	-	x	-	x	System	iterative	Service
Herrmann et al. [50]	-	x	-	x	x	Value Prop.	linear	Service
Bullinger et al. [51]	-	-	-	-	-	Value Prop.	iterative	Service
Kersten et al. [52]	-	x	x	-	x	Value Prop.	linear	Service
Lindahl et al. [53]	-	-	-	-	-	Value Prop.	linear	PSS
Botta [18]	x	x	x	-	-	Value Prop.	iterative	PSS
Tuli et al. [54]	-	-	x	-	x	Value Prop.	linear	PSS
PAS 1082 [55]	x	-	-	-	x	System	iterative	Service
Becker et al. [9]	-	-	-	-	-	Value Prop.	linear	PSS
Vasantha et al. [56]	-	-	-	-	x	System	iterative	PSS
Kim, et al. [57]	x	-	x	x	-	Value Prop.	linear	PSS
Müller [45]	x	x	x	-	x	System	linear	PSS

Our literature analysis of (product-)service (systems) engineering methods revealed four insights. First, few methods take a service system perspective, but rather focus on engineering a value proposition. Only ten methods consider the resources of customers, but limit the customers' role to requirements or need elicitation.

Second, using the identified basic operations of recombinant innovations as device of mind reveals that addition, dissociation, and association are seldom included in available service engineering methods. Twelve of the 24 analyzed approaches do not cover one of the stated operations at all, including all considered NSD approaches. Although eleven of the twelve remaining approaches include association, only four methods feature the operation transfer, which shows the largest gap.

Third, value propositions are not always perceived as solutions that can comprise both physical goods and services. Many approaches refer merely to engineering services without reference to any physical goods. Although only eight of the evaluated approaches focus on combining physical goods and services into customer solutions, these approaches became more frequent recently. Since 2006 all developed methods except one target PSS, which reveals a clear trend towards introducing all available types of resources into the co-creation of value.

Fourth, all evaluated approaches represent sequential or iterative processes and do not feature a prototypical approach as it is much discussed in Design Thinking or Software Engineering nowadays. As product development models are often linear approaches, some adapted methods for service engineering retained this structure.

5 Discussion

5.1 Design Principles for Service System Engineering Methods

The design principles are based on the four insights from literature analysis. They can be regarded as theory for design of service system engineering methods [58]. Thereby, the design principles explicitly prescribe how to build a service system engineering approach for recombinant innovation.

Design Principle 1: Take a service system engineering perspective

The analysis reveals that many service engineering methods present processes for designing a value proposition that is offered to a customer, but they refrain from specifying how the co-creation of value would be performed. Many approaches seem to implicitly take a goods-dominant logic perspective [4], [6] in which "services", "customer solutions", or "product-service systems" are engineered as marketable solutions, while refraining to specify the properties of a service system as a socio-technical system. This perspective is in line with methods for product engineering, foremost with the VDI-Standard 2221 [59], according to which many service engineering methods were designed. Even in Service Science, early papers defined a product-service system as "a mix of tangible products and intangible services designed and combined so that they jointly are capable of fulfilling final customer needs" [60]. As a result, "service engineering models, methods, and tools rarely focus on the development of service architectures" [2]. We argue that with the proliferation of technology in all societal sub-systems, integrating fragmented resources in socio-technical service systems will be increasingly crucial to provide superior value-in-use.

In line with this claim, the service-dominant logic of marketing [4], [6] posits that companies cannot offer *service* per se, as they can only offer *value propositions* that are enacted through a value co-creation of service providers and service customers, creating value-in-use for the actors involved. “Service systems comprise service providers and service clients working together to coproduce value in complex value chains or networks” [61]. Later, service systems were coined the basic abstraction in Service Science and defined as “a dynamic configuration of resources, including people, organizations, shared information (language, laws, measures, methods), and technology, all connected internally and externally to other service systems by value propositions“ [7], [61]. (Product-)service systems are socio-technical systems that enable co-creation of value by service providers and service customers [3], [2].

We argue that the design of service systems has to take a broader account than specifying the value proposition offered. But it should also focus the organizational and technological context that is required to turn a *value proposition* into *value-in-use*. Organizational and technological context comprise the assets and core competences that are brought to bear on the co-creation of value by (networks of) service providers and (networks of) service customers, including people, assets, processes, information systems and data, money, and relations with other actors that participate in a service (eco-)system. This view is beyond an abstract ‘implementation’ phase—as included in many current methods—since implementation refers to building up internal resources.

Design Principle 2: Recombine related resources in service systems
The analysis reveals that few of the reviewed service engineering methods refer to all aspects that constitute recombinant (service) innovation. Instead, many methods seem to perceive service engineering as a genuinely creative process that is conducted to design new value propositions from scratch, while not explicitly reusing or integrating solutions that are available in the service (eco-)system. As opposed to this finding, Brynjolfsson and McAfee [10] argue that in our *Second Machine Age* most innovations will be created based on recombining existing resources.

Recombinant innovation differs from configuring value propositions based on predefined modules. Methods for service modularization and configuration [35], [36], [62] usually presuppose that a finite solution space can be designed that is constrained by all permutations of the specified modules. Müller [45] refers to this approach as the *configuration shortcut*. In contrast, our approach is focused on service engineering itself, in which solutions are recombined to identify any missing modules that are required to set up a new value proposition and to co-create value-in-use.

In line with addition, dissociation, association, and transfer as the basic operations of recombinant innovation, we argue that service system engineering methods should explicitly identify the properties of the current service system as well as the value-propositions that can be designed and co-created with these resources. This relational approach goes beyond many available approaches [63], most of which focus on requirements elicitation and analysis. We argue that this perspective is inherently goods-dominant, since it does not put assets and core competencies of the involved stakeholders center-stage. As a result, requirements analysis is often not described as relational process, but as activity that is performed by a service provider alone.

Based on the socio-technical properties of service systems, a relational view on service system engineering would fit assets and core competencies of the involved actors together (association), further advance and detail existing assets and core competencies (dissociation), engineer new value propositions and transform the service system (addition), and apply resources outside their intended context of use (transfer).

Design Principle 3: Conceptualize value-in-use as based on both access to external resources and transfer of ownership of physical goods

Not surprisingly, our analysis refers to different types of service engineering approaches. While many methods in NSD focus on designing immaterial value-propositions, Service Engineering often explicitly integrates physical goods and services. Service research in the latter stream has come a long way from hybrid value-creation [64] to cyber-physical systems that view smart objects as resources in service systems.

Since we expect that many future service systems will be based on data and functionality provided by smart objects, we strongly argue that service systems must be designed to explicitly consider all resources that are available for recombination. Supporting this view, service-dominant logic [4], [6] has long advocated that physical goods are distribution mechanisms for service, since they stem from operand resources that stakeholders contribute to service systems. The rental-access paradigm [65] has highlighted that the core of service (as opposed to transferring ownership of products) is temporary access to resources, which customers do not own themselves.

Design Principle 4: Use prototypes to communicate the design of service systems

The analysis revealed that many methods conceptualize service engineering as a linear or iterative process, but seldom suggest prototypical approaches. Service engineering methods, in particular, feature many steps before a value proposition is offered to a customer. In contrast, innovation literature states that innovation is emergent and can even happen unintentionally or unplanned [24].

Recently, the Design Thinking movement has argued strongly for organizing engineering as a cyclic process that comprises several design cycles, each of which ends with a prototype. Similar approaches have been applied in software engineering for some time, including Scrum and other agile software development methods. A cyclic approach is also in line with the basic tenets of design science research that conceptualize design as building and evaluation [66], or with the cyclic approach presented in Action Design Research [67].

Since future service systems will rely strongly on data and information systems, we propose that they need to be designed with agile methods, as those proposed in software engineering. Service system engineering methods should, therefore, feature cycles of design and evaluation, followed by processes of organizational learning.

5.2 A Service Systems Engineering Approach for Recombinant Innovation

We visualize the design principles by presenting a conceptual method for service system engineering (Figure 2). The method builds on many established concepts of service engineering, since we reviewed all steps that current service engineering methods feature, and reorganized them for compatibility with our design principles. Our method comprises Service System Analysis, Service System Design, and Service System Transformation as its three basic sub-processes.

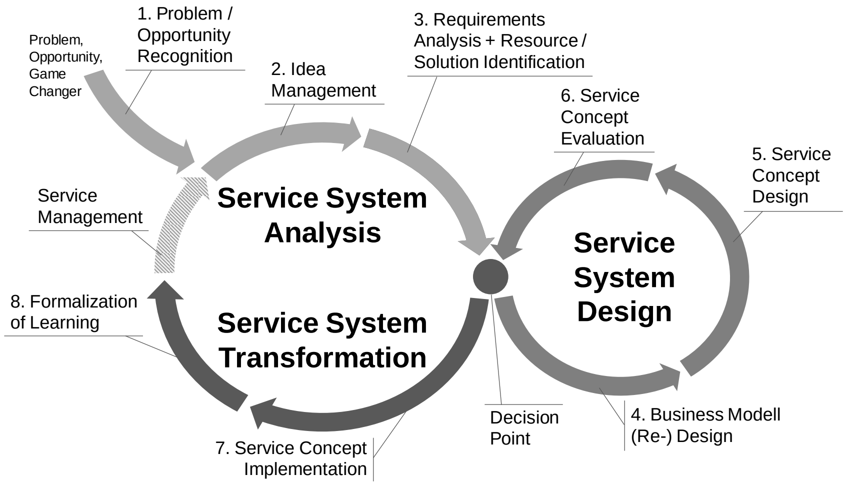


Figure 2: A method for recombinant service system engineering

Service System Analysis is started to remedy a problem or to realize an opportunity by (re-)designing a service system. In Idea Management, different ideas are identified and evaluated to identify those that are worth pursuing in a detailed analysis. Subsequently, an extensive Requirements Analysis is performed. Extending current methods, Requirements Analysis explicitly identifies the resources present in a service system, to enable the involved actors to recombine their assets and core competencies.

Service System Design comprises Business Model Design, followed by Service Concept Design, and Service Concept Evaluation. These activities are organized is cycles, in line with the design science paradigm that conceptualizes design as “to build” and “to evaluate” [66]. The Service Concept Design can comprise physical goods whose ownership is transferred in exchange processes and resources that can be rented/accessed by other actors in a service system.

Service System Transformation comprises implementing the final Service System Concept in order to integrate further resources and learn additional core competencies that are required to co-create the intended value-in-use. Therefore, the service system is transformed as a socio-technical system, beyond designing value propositions and then implementing resources. The service system engineering process is dynamic and focuses on developing viable prototypes as result of each cycle.

A **Decision Point** connects all three sub-processes. After Requirements Analysis is completed, service system engineers can decide to either recombine existing resources (transfer, association) and commence with Service Concept Implementation, or to commence with Service Concept Design (addition, dissociation). At the same time, the decision point marks the judgement to leave the design cycle after a viable Service Concept has been designed that complies with the identified requirements.

6 Conclusion

Based on a conceptual analysis of service engineering methods, our paper offers three contributions to research and practice. First, we provide an update on the methods for service engineering that have been proposed in various research streams in the Service Science discipline. Second, we identified conceptual shortcomings with respect to (a) applying a service systems perspective, (b) considering the basic operations of recombinant innovation that will likely become more prominent in what Brynjolfsson and McAfee [10] have termed the Second Machine Age, (c) considering the transfer of ownership for physical goods and the rental of/access to external resources, and (d) prototypes for designing value propositions. Third, we proposed design principles for service system engineering and visualized them with a conceptual model.

Limitations refer to the lack of generalizability that is inherent to conceptual and qualitative research. While we took precautions to objectify the coding process and attain inter-coder reliability, we acknowledge that other researchers might have come to different assessments of the reviewed service engineering methods.

Other researchers and practitioners can build on our results in multiple ways. First, as an IT artifact the design principles can be subjected to demonstration and evaluation that inform further design cycles. In particular, we are eager to see how the four basic operations of recombinant innovation can be applied successfully. Second, an evaluation could also shed light on how intensively or loosely product engineering and service systems engineering should be intertwined. While a close integration seems favorable to design consistent value propositions, loosely coupling the processes could keep the design of service systems agile, while decoupling them from more inflexible product development processes. Third, researches should investigate if organization do have the necessary resources at their disposal to implement the proposed approach. Concomitant, investigating synergist effects can be another area for future research.

References

1. Cavalieri, S., Pezzotta, G.: Product-Service Systems Engineering: State of the Art and Research Challenges. Computers in Industry 63, 278–288 (2012)
2. Böhmann, T., Leimeister, J.M., Möslein, K.: Service Systems Engineering. A Field for Future Information Systems Research. Business & Information Systems Engineering 6, 73–79 (2014)
3. Becker, J., Beverungen, D.F., Knackstedt, R.: The challenge of conceptual modeling for product–service systems: status-quo and perspectives for reference models and modeling languages. Information Systems and e-Business Management 8, 33–66 (2009)
4. Vargo, S.L., Lusch, R.F.: Why "service"? Journal of the Academy of Marketing Science 36, 25–38 (2008)
5. Fähnrich, K.-P., Opitz, M.: Service Engineering - Entwicklungspfad und Bild einer jungen Disziplin. In: Bullinger, H.-J., Scheer, A.-W. (eds.) Service Engineering, pp. 85–112. Springer, Berlin (2006)

6. Vargo, S.L., Lusch, R.F.: Service-Dominant Logic: Continuing the Evolution. Journal of the Academy of Marketing Science 36, 1–10 (2008)
7. Maglio, P.P., Vargo, S.L., Caswell, N., Spohrer, J.: The Service System is the Basic Abstraction of Service Science. Information Systems and e-Business Management 7, 395–406 (2009)
8. Meyer, K., Böttcher, M.: Entwicklungspfad Service Engineering 2.0. Neue Perspektiven für die Dienstleistungsentwicklung. Univ, Leipzig (2011)
9. Becker, J., Beverungen, D., Matzner, M., Müller, O., Pöppelbuß, J.: Design Science in Service Research: A Framework-Based Review of IT Artifacts in Germany. In: Jain, H., Sinha, A.P., Vitharana, P. (eds.) Proceedings of the Sixth International Conference on Design Science Research in Information Systems and Technology (DESRIST 2011), 6629, pp. 366–375. Springer, Berlin (2011)
10. Brynjolfsson, E., McAfee, A.: Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. Norton & Company, New York City, NY, USA (2014)
11. Meiren, T., Barth, T.: Service Engineering in Unternehmen umsetzen. Leitfaden für die Entwicklung von Dienstleistungen. Fraunhofer-IRB-Verl., Stuttgart (2002)
12. Johnson, S.P., Menor, L.J., Roth, A.V., Chase, R.B.: A Critical Evaluation of the New Service Development Process. In: Fitzsimmons, J.A., Fitzsimmons Mona J. (eds.) New Service Development, pp. 1–32. SAGE Publications (2000)
13. Edvardsson, B., Olsson, J.: Key Concepts for New Service Development. The Service Industries Journal 16, 140–164 (1996)
14. Ramaswamy, R.: Design and Management of Service Processes. Keeping customers for life. Addison-Wesley, Reading, Massachusetts (1996)
15. Shostack, L.G., Kingman-Brundage, J.: How to Design a Service. In: Congram, C.A., Friedman Margaret L. (eds.) The AMA Handbook of Marketing for the Service Industries, pp. 243–261 (1991)
16. Scheuing, E.E., Johnson, E.M.: A Proposed Model for New Service Development. Journal of Services Marketing 3, 25–34 (1989)
17. Klein, R.: Modellgestütztes Service Systems Engineering. Theorie und Technik einer systemischen Entwicklung von Dienstleistungen. Dt. Univ.-Verl., Wiesbaden (2007)
18. Botta, C.: Rahmenkonzept zur Entwicklung von Product-Service Systems. Product-Service Systems Engineering. Josef Eul Verlag, Lohmar, Köln (2007)
19. Jaschinski, C.: Qualitätsorientiertes Redesign von Dienstleistungen. Shaker, Aachen (1998)
20. DIN Deutsches Institut für Normung e. V.: DIN-Fachbericht 75 - Service Engineering. Beuth Verlag GmbH, Berlin (2008)
21. Toivonen, M., Tuominen, T.: Emergence of Innovations in Services. The Service Industries Journal 29, 887–902 (2009)
22. Cooke, P.: Four Minutes to Four Years. The Advantage of Recombinant over Specialized Innovation – RIS3 versus 'Smartspec'. In: European Planning Studies, 24, pp. 1494–1510
23. Gallouj, F., Weinstein, O.: Innovation in services. Research Policy 26, 537 (1997)
24. Gremyr, I., Witell, L., Löfberg, N., Edvardsson, B., Fundin, A.: Understanding New Service Development and Service Innovation through Innovation Modes. Journal of Business & Industrial Marketing 29, 123–131 (2014)
25. Wirth, M., Friesike, S., Flath, C., Thiesse, F.: Patterns of Remixes or Where Do Innovations Come From: Evidence from 3D Printing. In: Proceedings of the 23rd European Conference on Information Systems (ECIS), Münster (2015)
26. Fleming, L.: Recombinant Uncertainty in Technological Search. Management Science 47, 117 (2001)

27. Cecere, G., Ozman, M.: Innovation, Recombination and Technological Proximity. Journal of the Knowledge Economy 5, 646–667 (2014)
28. Kim, J., Park, Y.: E-Service Concept Design in Recombinative Innovation: A Morphology Analysis Approach. In: Proceedings of the 19th International Conference on Industrial Engineering and Engineering Management (IEEM), Hong Kong (2012), pp. 666–670
29. Tsur, Y., Zemel, A.: Towards Endogenous Recombinant Growth. Journal of Economic Dynamics and Control 31, 3459–3477 (2007)
30. Gadrey, J., Gallouj, F., Weinstein, O.: New Modes of Innovation. International Journal of Service Industry Management 6, 4–16 (1995)
31. Antonelli, C., Krafft, J., Quatraro, F.: Recombinant knowledge and growth. The case of ICTs. Structural Change and Economic Dynamics 21, 50–69 (2010)
32. Windahl, C., Andersson, P., Berggren, C., Nehler, C.: Manufacturing Firms and Integrated Solutions. European Journal of Innovation Management 7, 218–228 (2004)
33. Gilmore, J.H., Pine II, B.J.: The four faces of mass customization. Harvard Business Review 75, 91–101 (1997)
34. Böhmann, T., Langer, P., Schermann, M.: Systematische Überführung von kundenspezifischen IT-Lösungen in integrierte Produkt-Dienstleistungsbausteine mit der SCORE-Methode. WIRTSCHAFTSINFORMATIK 50, 196–207 (2008)
35. Becker, J., Beverungen, D., Knackstedt, R., Müller, O.: Pricing of Value Bundles: A Multi-Perspective Decision Support Approach. Enterprise Modelling and Information Systems Architectures (EMISA) 6, 54–69 (2011)
36. Becker, J., Beverungen, D., Knackstedt, R., Müller, O.: Model-Based Decision Support for the Customer-Specific Configuration of Value Bundles. Enterprise Modelling and Information Systems Architectures (EMISA) 4, 26–38 (2009)
37. Dörbecker, R., Boehmann, T.: Tackling the Granularity Problem in Service Modularization. In: Proceedings of the 21st American Conference on Information Systems (AMCIS), Puerto Rico (2015)
38. Baida, Z.: Software-Aided Service Bundling: Intelligent Methods and Tools for Graphical Service Modelling (2006)
39. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a literature Review. MIS Quarterly 26, 13–23 (2002)
40. Krippendorff, K.: Content Analysis: An Introduction to Its Methodology. SAGE Publications, Thousand Oaks, CA, USA (2013)
41. Yadav, M.S.: The Decline of Conceptual Articles and Implications for Knowledge Development. Journal of Marketing 74, 1–19 (2010)
42. Senyard, J., Baker, T., Steffens, P., Davidsson, P.: Bricolage as a Path to Innovativeness for Resource-Constrained New Firms. Journal of Product Innovation Management 31, 211–230 (2014)
43. Schneider, K., Scheer, A.-W.: Konzept zur systematischen und kundenorientierten Entwicklng von Dienstleistungen. Veröffentlichungen des Instituts für Wirtschaftsinformatik 175, 1–45 (2003)
44. Kolko, J.: Design Thinking Comes of Age. Harvard Business Review, 66–71 (2015)
45. Müller, P.: Integrated Engineering of Products and Services. Fraunhofer, Stuttgart (2014)
46. Schwarz, W.: Methodisches Konstruieren als Mittel zur systematischen Gestaltung von Dienstleistungen. IPK, Berlin (1997)
47. Schreiner, P., Klein, L., Seemann, C.: Die Dienstleistungen im Griff - erfolgreich gründen mit System. Fraunhofer-IRB-Verl., Stuttgart (2001)
48. Morelli, N.: Product-service systems, a perspective shift for designers. A case study: the design of a telecentre. Design Studies 24, 73–99 (2003)

49. Kunau, G., Junginger, M., Herrmann, T., Krcmar, H.: Ein Referenzmodell für das Service Engineering mit multiperspektivischem Ansatz. In: Herrmann, T. (ed.) Konzepte für das Service-Engineering, pp. 187–216. Physica-Verl., Heidelberg (2005)
50. Herrmann, K., Klein, R., The, T.-S.: Computer Aided Service Engineering-Konzeption eines Service Engineering Tools. In: Bullinger, H.-J., Scheer, A.-W. (eds.) Service Engineering, pp. 649–678. Springer, Berlin (2006)
51. Bullinger, H.-J., Schreiner, P.: Service Engineering: Ein Rahmenkonzept für die systematische Entwicklung von Dienstleistungen. In: Bullinger, H.-J., Scheer, A.-W. (eds.) Service Engineering, pp. 53–84. Springer, Berlin (2006)
52. Kersten, W., Kern, E.-M., Zink, T.: Collaborative Service Engineering. In: Bullinger, H.-J., Scheer, A.-W. (eds.) Service Engineering, pp. 341–357. Springer, Berlin (2006)
53. Lindahl, M., Sundin, E., Sakao, T., Shimomura, Y.: An interactive design methodology for service engineering of functional sales concepts: a potential design for environment methodology. In: Proceedings of the 13th CIRP International Conference on Life Cycle Engineering (LCE), pp. 589–594. Leuven, Belgium (2006)
54. Tuli, K.R., Kohli, A.K., Bharadwaj, S.G.: Rethinking Customer Solutions: From Product Bundles to Relational Processes. Journal of Marketing 71, 1–17 (2007)
55. German Standards Institute: Publicly Available Specification (PAS) 1082: Standardisierter Prozess zur Entwicklung industrieller Dienstleistungen in Netzwerken. Beuth, Berlin (2008)
56. Vasantha, A., Vijaykumar, G., Hussain, R., Roy, R., Tiwari, A., Evans, S.: A Framework for Designing Product-Service Systems. In: Proceedings of the 18th International Conference on Engineering Design (ICED11), Copenhagen, Denmark (2011)
57. Kim, K.-J., Lim, C.-H., Lee, D.-H., Lee, J., Hong, Y.-S., Park, K.: A Concept Generation Support System for Product-Service System Development. Service Science 4, 349–364 (2012)
58. Chandra, L., Seidel, S., Gregor, S.: Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions. In: 2015 48th Hawaii International Conference on System Sciences (HICSS), pp. 4039–4048
59. Verein Deutscher Ingenieure: Methodik zum Entwickeln und Konstruieren technischer Systeme und Produkte. Beuth Verlag, Berlin (1993)
60. Tukker, A., Tischner, U.: Product-Services as a Research Field: Past, Present and Future. Reflections from a Decade of Research. Journal of cleaner production 14, 1552–1556 (2006)
61. Spohrer, J., Maglio, P.P., Bailey, J., Gruhl, D.: Steps Toward a Science of Service Systems. IEEE Computer 40, 71–77 (2007)
62. Dörbecker, R., Böhmann, T.: FAMouS - Framework for Architecting Modular Services. In: Proceedings of the Thirty Sixth International Conference on Information Systems (ICIS), pp. 1–18. Fort Worth (2015)
63. Dyer, J.H., Singh, H.: The Relational View: Cooperative Strategy and Sources of Inter-organizational Competitive Advantage. Academy of Management Review 23, 660–679 (1998)
64. German Standards Institute: Publicly Available Specification (PAS) 1094: Product-Service Systems – Value Creation by Integrating Goods and Services. Beuth, Berlin (2009)
65. Lovelock, C., Gummesson, E.: Whither Services Marketing?: In Search of a New Paradigm and Fresh Perspectives. Journal of Service Research 7, 20–41 (2004)
66. March, S.T., Smith, G.F.: Design and Natural Science Research on Information Technology. Decision Support Systems 15, 251–266 (1995)
67. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R., Saunders, C., Iivari, J.: Action Design Research. Management Information Systems Quarterly 35, 37–56 (2011)

Blueprinting Crowdfunding
Designing a Crowdfunding Service Configuration Framework

Philipp Haas, Ivo Blohm

University of St.Gallen, Institute of Information Management, St. Gallen, Switzerland
{philipp.haas,ivo.blohm}@unisg.ch

Abstract. Crowdfunding gained momentum over the last few years. In contrast to traditional forms of funding, the service provision of crowdfunding platforms is performed within service systems. These comprise a complex combination of IT and non-IT services, different stakeholders, and diverging contexts and purposes. The design and operation of such service systems represents a tough challenge. Therefore, we developed a crowdfunding service configuration framework in the form of a morphological box and derived three dominant design patterns by following a design science approach. Therefore, we followed three iterations, which comprise in total twelve expert interviews, three case studies and the analysis of 161 crowdfunding platforms. The configuration framework extends research on crowdfunding and service science by providing insights in how to support the systematic design of crowdfunding service systems, reducing their complexity, and giving a comprehensive overview over their building blocks.

Keywords: Crowdfunding, Service Systems, Modularization, Design Science

1 Introduction

Crowdfunding represents a new way of funding projects or companies, involving a diverse crowd of private capital givers over the Internet, and is frequently considered a more transparent, easy, entertaining, and democratic way of funding. Therefore, crowdfunding gained momentum during the last few years and began to establish as an alternative way of funding. As a consequence a variety of complementary crowdfunding platforms emerged, ranging from altruistic to profit oriented offerings. These mostly start-up driven crowdfunding platform providers build innovative offerings for both, the utilization of highly specialized niche markets as well as the mass market for financial products. Crowdfunding start-ups use their high degree of automation, the Internet, the web 2.0, and innovative opportunities such as data analytics. In order to grasp these opportunities, the service provision of crowdfunding platforms is performed within service systems. This allows operators of crowdfunding platforms to provide some services by themselves, whereas they may source others from specialized partners (e.g. payment, banking, dunning) within a service system [1].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Haas, P.; Blohm, I. (2017): Blueprinting Crowdfunding - Designing a Crowdfunding Service Configuration Framework, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 151-165

Despite the huge growth of the crowdfunding market in terms of origination volume and platform numbers, this growth is not distributed equally among all types of crowdfunding. While the market for crowdlending is booming, the market for crowdinvesting is stagnating. Further, the market in general is characterized by a large fluctuation and shows the tendency of consolidation [2, 3]. As crowdfunding comprises a complex combination of services and stakeholders, the design of such service systems represents a tough challenge. Thus, many attempts to design new crowdfunding service systems struggle, as the complexity of the crowdfunding service system can't be overseen and it lacks knowledge about how to systematically design crowdfunding service systems. In order to overcome this challenge the design has to allow the decomposition of the crowdfunding service system into single components. This approach is known from the concept of service modularization [4, 5]. Especially, during early stages of the development of crowdfunding services and the assessment of design choices, guidance is needed. Despite its relevance, research on crowdfunding has largely neglected the topics of how to systematically design crowdfunding service systems [6]. This hampers the development of new crowdfunding offerings, the exploitations of new market niches and the maturation of the crowdfunding industry. Therefore, this paper pursues the research question of *how potential crowdfunding providers can design crowdfunding service systems systematically*.

In order to answer this research question, this paper follows the design science (DS) paradigm [7, 8] by designing a crowdfunding service configuration framework, which takes the form of a morphological box, by combining a component perspective and a functional perspective. The crowdfunding service configuration framework aims at a structured and comprehensive presentation of crowdfunding service systems by presenting ten constituting service modules with in total 24 differentiating parameters and three dominant design patterns – altruism, hedonism, and profit-orientation.

This paper provides especially two theoretical contributions. First, the paper expands crowdfunding research by proposing a crowdfunding service configuration framework, which describes the building modules of a crowdfunding service system. Thus, the configuration framework provides an overview over required service modules and respective parameters. Second, the paper bridges research on crowdfunding with the field of service science. Thus, we contribute to service science, by proposing a framework for the systematic design of modularized services systems. Therefore, we contribute to the call for the design of novel artefacts, facilitating the engineering and management of service systems [4]. For practice, this paper provides guidance for the systematic design of crowdfunding service systems, the decision support for the assessment of required competences, the identification of white spots for business opportunities, and a better understanding of the disruptive potential of crowdfunding.

The paper proceeds as follows. First, we give an overview over the related work regarding complex service systems and crowdfunding service systems. Second, our DSR approach is presented. Third, the iterative design process of the crowdfunding service configuration framework is described. Fourth, we discuss our findings and present our theoretical and practical contributions.

2 Related Work

2.1 Complex Service Systems

A service is a bundle of activities, which takes place between the user and provider of a service within a service system [9, 10]. Service systems can be described as value-co-creation of stakeholders, technologies, and shared information (e.g., language, laws, measures, and methods) [11]. One can speak of a complex service, if multiple stakeholders are integrated into the service system, various interactions between them are necessary, and IT and non-IT activities are combined in order to harness its value [12]. In order to leverage the value co-creation, service systems follow a modular design, which enables the systematic engineering of service systems [4, 13]. Modularization can be described as decomposition of a single object into decoupled single components, which can be combined in various way to create new configurations [14]. Modularization rests upon the basic principles of cohesion and loose coupling [15, 16] and has been already established in the context of service science [17, 18].

2.2 Crowdfunding Service Systems

Analogous to crowdsourcing, crowdfunding can be defined as collective financing by an undefined crowd by means of an internet-based open call [19, 20]. Despite the large attention the topic of crowdfunding has drawn among the financial service industry, research on crowdfunding has largely neglected the systematic design of crowdfunding service systems. Certain studies aimed at the systemization of crowdfunding services systems [20-23]. Tomczak and Brem [24] aimed at conceptualizing an investment model by taking a process perspective on crowdfunding service systems by applying process modeling technique. Wieck et al. [25] made an early attempt to investigate how to develop, pilot and evaluate an crowdfunding service system, in order to support university startups. Liebenau et al. [26], Hemer [27], and O`Sullivan et al. [28] argue that the advantage of crowdfunding service systems lies in their modular ecosystem structure, which enables the bundling and aggregation of various competences within a complex service system. Recently, Haas et al. [6] reported about the implementation of the modular design of a crowdfunding service system.

3 Research Approach

In order to develop a configuration framework for the systematic design of crowdfunding service systems, a morphological box turned out to be a valid form. A morphological box combines a component perspective by listing the building modules and functional perspective by detailing these modules in single functional characteristics. This makes a morphological box a heuristic method for capturing complex issues [29, 30] such as the design of crowdfunding service systems. Besides, morphological boxes have been successfully applied to the context of complex IT services before [31, 32].

In order to develop and evaluate the crowdfunding service configuration framework, which supports potential providers of crowdfunding service system, we followed a design science approach. Design science research is highly suitable in solving a real world problem such as the systematic design of crowdfunding service systems. Therefore, design science aims at the iterative development of an innovative IT artefact. The design science paradigm, as suggested by Hevner [7, 8], aims at rigor and relevance of the proposed design by following three integrated cycles: relevance cycle, design cycle, and rigor cycle. The relevance cycle aims at bridging the design activities with its practical environment. Thus, it helps specifying the real-world problem, eliciting the needs and requirements for solving it, and the recirculation of the designed artefact to the field of practice. The rigor cycle ensures the interconnection between the designing of the artefact with the existing knowledge base. Thus, the design of the artefact is informed by existing theories and knowledge, while new knowledge, resulting from the design, is recirculated to the knowledge base. Surrounded and influenced by the relevance cycle and the rigor cycle, the design circle is situated in the center. The design cycle represents the iterative design activities, which are necessary in order to construct and evaluate the artefact.

We performed three iterations in order to design and evaluate the crowdfunding service system configuration framework. Within the first design iteration, the problem has been specified and requirements have been elicited from the field. Therefore, we conducted a comprehensive study of related literature and performed three expert interviews. All three interviewed experts aimed at engaging in the crowdfunding market by designing an own crowdfunding service system, but struggled, as they were unable to oversee the complexity of the crowdfunding service systems. The interviews led to a first impression of the scope and form of the configuration framework. Within a second phase, three case studies, comprising one illustrative example for each crowdfunding service system archetype – altruistic, hedonistic, and profit-oriented [22] – have been conducted in order to identify the building components and characteristics of crowdfunding service systems. Further, service modularization technique has been applied [16, 33]. Therefore, the identified services within the crowdfunding service system have been described on a process level in order to derive modularization parameters. Afterwards, these parameters are applied in order to identify the actual service modules. Thereby, a first version of the configuration framework has been designed. In order to evaluate the module validity, the framework's comprehensiveness, its applicability, and usefulness, six interviews have been conducted with crowdfunding experts (bank representatives, platform providers, and researchers), which participated in the design of a crowdfunding service system. This led to further refinement of the artefact. Within a final phase, the configuration framework has been applied to code 161 crowdfunding service systems, in order to identify dominant design patterns, which could serve as starting point for the design of crowdfunding service systems. A final evaluation of the configuration framework and the patterns has been conducted by interviewing the three initial experts again. The experts have been questioned whether the configuration framework meets their mentioned requirements and, by looking back, whether the identified design patterns represent suitable starting points for the design of their crowdfunding service systems.

4 The Crowdfunding Service Configuration Framework

4.1 Iteration 1: Problem Specification and Requirements

Within the first design iteration, including all three cycles, the aim was to specify the problem and to elicit requirements for the configuration framework. In order to specify the problem and elicit requirements from the field, three expert interviews and a comprehensive literature study have been conducted. First, we started performing a relevance cycle by conducting three expert interviews in order to get an impression of the problem of designing crowdfunding service systems and in order to elicit design requirements. The interviews were conducted via Skype during June 2016 and were 30 to 60 minutes long. The interviewees came from two different banks and a start-up incubator. All three experts were responsible for the design of crowdfunding service systems in distinguishing contexts. All three struggled with their attempts to engage in the crowdfunding market, as they overstrained with the complexity of the crowdfunding service systems. They annotated consistently that especially during the beginning of their attempts, they longed for support in overseeing alternative options and dependencies. They had to waste a lot of time and resources in order to figure out basic functionalities of value proposition, value creation, and value capturing and assessing the general fit of a crowdfunding type to their desired objectives. The input from the relevance cycle has been expanded by performing a rigor cycle. Therefore, findings from a comprehensive literature study, regarding literature on crowdfunding and complex service systems has been used to inform the elicitation of the requirements and to bridge the different literature streams, in order to enhance the current body of knowledge. After finishing the rigor cycle, we evaluated and refined our recent design activities – the deducted requirements – by interviewing our experts again, in order to ensure comprehensibility, correctness and applicability. Iteration 1 identified three major requirements: 1) Early-stage applicability and reduction of complexity. 2) Structured and comprehensive presentation of crowdfunding service systems. 3) Dominant design patterns as template. Table 1 gives an overview over the identified and evaluated requirements.

Table 1: *Design Requirements*

Requirement	*Description*
Early-stage applicability and reduction of complexity	Crowdfunding is based on components and competencies, which have not been considered relevant so far [6, 26]. Thus, many struggle at early design stages to oversee its complexity and disruptive potential and lack critical competencies. Many different stakeholders are necessary in order to bundle the required knowledge. Therefore, complexity has to be reduced in a heuristic manner, in order to light up the opportunities, objectives, functionalities, and consequences of crowdfunding for the involved stakeholders.

Structured and comprehensive presentation of crowdfunding service systems	Due to its high complexity, the various functionalities and dependencies within the service system are hard to oversee. Therefore, a functional perspective, as well as a component perspective, have to be combined in order to structure the constituting components of a crowdfunding service system. Besides the comprehensive overview over the single services, ensuring flexibility for several configurations is paramount. Therefore, a modular structure of the implicated crowdfunding services within the framework enables the loose coupling and thereby, easy reconfiguration of the components.
Dominant design patterns as template	As crowdfunding service systems can be designed for various purposes, the definition of what to achieve with an own crowdfunding service system and which configuration supports these objectives is often blurry. Providing dominant design patterns have to be identified in order to serve as a starting point for the design activities.

Besides the deducted requirements, Iteration 1 led to a first impression of the scope of the configuration framework and identified a morphological box as a valid and suitable form, due to its ability to capture complex issues, bridging a functional and a component perspective. Its heuristic character reduces complexity and enables early-stage application even for unexperienced co-workers.

4.2 Iteration 2: Designing the Configuration Framework

After specifying the problem and eliciting requirements from the field, we conducted three case studies of the three experts' initiatives for designing a crowdfunding service system, in order to identify the building modules of crowdfunding service systems. These cases represent illustrative examples for each archetype of crowdfunding service system - altruistic, hedonistic, and profit-oriented [22]. In order to collect the data for the case studies and evaluate our findings, we conducted multiple iterative interviews and workshops with the experts and the respective members of the project teams. Further, we studied the business models of each case example by analyzing public information (e.g., website, terms & conditions) and private documents (e.g. business plans, process models). In order to perform a rigor cycle, we studied literature regarding process and ecosystem modelling and service modularization in order to find heuristic methods for their illustrations and analyzes. Thus, we identified three suitable methods - activity chain modelling for processes [34], e3 value for the illustration of ecosystems [35], and TM3 as method for service modularization [16]. We began the design cycle by modelling the customer journey and the ecosystems. We continued by modelling the single complementary activities of each stakeholder, which supports the customer journey or the crowdfunding process. Thereby, participating stakeholders, interfaces, information-, and money flows have been considered and evaluated. Afterwards, the activities have been modularized according to defined modularizing parameters, which aim at ensuring internal cohesion and loose coupling [16]. These parameters have been defined as 1) representing a pivotal topic within the crowdfunding process; 2) represents a closed activity; 3) is performed and provided by one stakeholder. Each identified service module represents a bundle of activities regarding specific processes within the configuration framework. These activities have been grouped by analyzing the intra service module cohesion, in order to identify the major parameters of a service module. The three case studies indicated a robust set of the similar ten service modules with in total 24 differentiating parameters. As the characteristic of the modules differentiate between each of the analyzed service systems, variations of the parameters have been defined. Defining these characteristics as variations of the module, allows the parallel selection of different characteristics for each module, within the crowdfunding service system. In order to ensure completeness and generalizability of our findings we performed another relevance cycle. We extended the identification of further parameter variations to a dataset of 161 crowdfunding service systems, which have been identified by conducting an online search. Search criterions included that: 1) it is active; 2) it is in German or English language; 3) the necessary information are publicly available; 4) it refers to a crowdfunding mechanism (e.g. mentioning the term crowdfunding). We reviewed each module parameter on each of the 161 crowdfunding service systems, included new variations and aggregated similar ones. In total one to six parameter variations have been identified and finally included in the crowdfunding service configuration framework.

Table 2 gives an overview over the identified service modules and the according characteristics.

Table 2: *Overview Service Modules*

Service Module	*Description*
Matchmaking	Matchmaking between capital givers and capital seekers represents a pivotal service within the service system. Therefore, an e-market place is operated in order to provide information, and to register funding decisions. As the matchmaking takes place in a two-sided market, we identified the two parameters capital seekers and capital givers, which showed two respectively three variations. Thus, we identified *individual* and *institutional capital givers* and *individuals, non-profit organizations, and for-profit organizations* as capital seekers.
Crowd Activation	Crowdfunding includes the attraction, activation, and balancing of the 'right' crowd of capital givers and seekers in order to ensure funding success, attractive returns, and to generate thick markets and network effects. Therefore, activating activities are performed online and offline. These two parameters showed three variations respectively – *none, mass advertising and personalized advertising*.
Customer Support	Crowdfunding aims at being more unbureaucratic and easier. Therefore, overcoming initial barriers and to clarify customer issues is addressed by providing comprehensive support for both capital givers and capital seekers. Both parameters showed the same five variations – *none, offline support, online support, personalized support, and automatized support.*

Market Differentiation	Crowdfunding mainly focuses on niche markets and serves the long tail of the financial service industry. Thus, it provides funding for project which cannot be served profitably by the traditional financial service industry. As crowdfunding service systems serve highly heterogeneous needs, a precise market differentiation is undertaken. Thus, we identified three market differentiating parameters – the motivation of the crowd, the market specialization of the service system, and the type of compensation, which is provided by the capital seekers. The motivation of the crowd differentiates between *altruism*, which aims at doing good, *hedonism*, which aims at satisfying own curiosity, and *profit-orientation*, which aims at satisfying monetary expectations. The specialization of crowdfunding intermediaries varies between *sustainability & social action, startup & new business, private consumption*, and *creative projects & products*. The compensations range from a *greater good*, where no compensation is provided, non-monetary *rewards*, *interest*, to proportional *profit-shares* according the success of the supported project.
Investor Relations	Crowdfunding as a more transparent and democratic way of investing aims at fostering communication between capital givers and capital seekers and enables a performance monitoring of the projects. The communication channels between capital givers and capital seekers revealed three variations- *none, traditional communication channels* (such as e-mails, telephone, fax), and *web 2.0 communication channels* (such as social media, blogs, and chats). As a second parameter performance monitoring is implemented by three variations – *none, progress bar*, which shows the actual funding status, or a *portfolio management system*, which enables an aggregated overview over the invested capital or even an automatized (re-)investment process regarding to the portfolio specifications.
Contracting	Contracting is essential for ensuring liability and compliance. Therefore, we identified two major parameters within this service module. First, terms and conditions mainly regulates the use of the crowdfunding service in general. We found four variations – *none, standardized terms of use, privacy policy regulations*, and *payment regulations*. Second, the legal relationship between capital seekers and capital givers after funding success represents a differentiating parameter. This parameter showed the two variations *direct legal relationship*, in the case of a direct peer-to-peer relationship, and *indirect relationship*, in the case of a legal intermediation (e.g., a bank).

Risk Assessment	Overcoming information asymmetries is essential in order to provide funding for capital seekers and reduce default risks for capital giver. Two parameters have been identified – due diligence and feasibility. The due diligence aims at assessing the credit-, and trustworthiness of a project and the capital seekers. The due diligence parameter shows three variations – *none, traditional forms,* by assessing personal data and documents, and *big data analyses*, which includes information based on data analytics (e.g., behavioral information). The second parameter aims at assessing the feasibility of a project, which can be performed by three variations – *none, business/project plan*, and *prototype.*
IT Functionality & Operations	A reliable platform with satisfying functionality is pivotal, as it represents the digital point of contact between capital seekers and givers. Overall, three parameters have been identified. First, the development and hosting of the platform, which shows the three variations *in-house, external service provision*, and *white-label solution.* Second, the registration process for capital givers and seekers, which is performed by the three variations *none, website login (via e-mail and password)*, or *social login* (Facebook or Google). Third, the form of the application can be differentiated into the two variations *web app*, or *mobile app*.
Payment	Payment represents a pivotal service as a fast, reliable, and efficient flow of money can be provided. Four parameters have been identified. First, the actual form of the payment system, which shows four variations – *offline payment* (e.g., cash in-payment), *traditional direct payment* (e.g., credit card), *online direct payment* (e.g., PayPal), and *direct debiting*. Second, the time of the payment, which can be *pre-paid, instant-paid*, and *post-paid*. Third, in case of debt default four variations can be differentiated – *none, notifications, dunning,* and *debt collection.* Fourth, the form of the payment processing – *directly* between the capital giver and seeker or *indirectly* via a financial intermediary (e.g., a bank).
Authentication	In order to meet certain legal regulations, prevent fraud, and reduce risks for capital seekers and givers, know your customer (KYC) services are applied regarding capital seekers and capital givers. Both parameters show four variations – *none, personal offline identification* (e.g., via a post office, notary), *automated digital identification* (e.g., digital passport, CAPTCHA), and personal online identification (e.g., via webcam)

These modules have been summarized within a morphological box, which represents the crowdfunding service configuration framework. For evaluating the proposed design of the configuration framework, with regard to module validity, the framework's comprehensiveness, its applicability, and usefulness, we conducted six interviews in total. Therefore, we re-interviewed the three initial experts plus three additional crowdfunding experts, which participated in the design or operation of a crowdfunding service system as well. One of the new consulted experts came from a bank and two from academia. First, the experts were asked to apply the configuration framework to their crowdfunding service system. Second, we asked them to rate the configuration framework with regard to comprehensiveness, its applicability, usefulness, and whether it meets the design requirements. The evaluation indicated good fit to the design requirements and confirmed comprehensiveness of the stated parameters and characteristics, high applicability for early design phases, and usefulness as it reduces complexity in a heuristic manner. The experts' feedback was taken into account thoroughly and led to further refinement of the configuration framework. The evaluated and refined version is presented in Table 3.

4.3 Iteration 3: Dominant Design Patterns

The three cases and our search for parameter variations revealed fundamental differences in the module characteristics and the module configurations. Nevertheless, we assumed the existence of dominant design patterns, as these differences are related to the basic orientation of the crowdfunding service system, which ranges from altruistic, hedonistic, to profit oriented purposes [22]. A rigor cycle regarding literature on the systemization of crowdfunding service systems revealed that these respective archetypes require different configurations due to differentiating target markets, related risks, legal reasons, and the motivation of capital givers and seekers [6, 21, 24]. Thus, the identification of basic design patterns would serve as a useful starting point for the design of crowdfunding service systems. Therefore, a relevance cycle has been conducted by applying the configuration framework to the 161 crowdfunding service systems from our previous platform analysis, which have been grouped according to its respective crowdfunding archetype – altruism (N=53), hedonism (N=60), and profit-orientation (N=48). Thus, the three groups showed large internal proximity with regard to four service modules, which differentiates clearly against the other groups - market differentiation, risk assessment, payment, and authentication. Thus, performing a design cycle, we defined three design patterns for crowdfunding service systems, which correspond to the three crowdfunding service system archetypes altruism, hedonism, and profit orientation. The predominant parameter variations of the three design patterns are indicated by color-coding in the configuration framework (see Table 3) - altruism: bright grey; hedonism: dark grey; profit-orientation: black.

The altruism design pattern is characterized by altruistic motives of the capital seekers and givers. Therefore, it focuses on sustainable and social caring projects and provides no compensation besides a sense for supporting a greater good. In contrast to the other design patterns non-governmental organizations appear as capital seekers. Typical examples for the altruistic design pattern might be *Benevolent*, *100Days*, or *Kiva*.

The hedonistic design pattern satisfies hedonistic motives and therefore, offers reward-based compensations and focuses mostly on the funding of creative projects. In order to reduce investment risks and to ensure the feasibility of the proposed crowdfunding projects, a feasibility check based on business or project plans or even prototypes is applied. Further, a basic level of activity in the case of debt default is performed by actively notifying defaulting capital seekers or givers. Typical examples for the hedonistic design pattern might be *Kickstarter*, *Startnext*, or *WeMakeIt*.

The most rigid pattern is represented by profit-oriented crowdfunding service systems, due to higher default risks and stronger legal regulation. Capital givers are motivated by gaining profits. Therefore, this pattern focuses on the funding of either start-ups or new businesses, where profit shares a predominant as compensation, or funding private consumption by granting loans and providing interests as compensation. Providing a portfolio-management system for fostering investor relations enables both, risk diversification and maximizing profits. Effective risk assessment is crucial due to the higher risk. Therefore, comprehensive due diligences based on traditional documentary are necessary. In the case of private capital seekers, these due diligences are often extended by data analyses based on the online behavior of the capital seekers (such as online times, previous visited websites, etc.). In the case of debt default, activities regarding dunning or even debt collection are predominant. Due to anti money laundering legislation, KYC activities are necessary in the profit-oriented design pattern. Typical examples might be *Companisto*, *Lendico*, or *Investiere*.

5 Discussion & Implications

This study presents a rigor and relevant crowdfunding service configuration framework in the form of a morphological box, which supports potential providers to systematically design crowdfunding service systems. By applying service modularization technique, we identified ten service modules, which represent required constituting blocks of a crowdfunding service system. This modules can be implemented via 24 module parameters with two to six parameter variations. Thus the parameter variations represent instantiations of a service module within a crowdfunding service system, which represents design choices for the early-stage blueprinting of crowdfunding service systems. Our evaluation showed that the configuration framework is comprehensive, useful, and applicable. Further, we derived three dominant design patterns – altruism, hedonism, and profit-orientation. Thus, this patterns support previous findings of crowdfunding research [22]. We identified strong in-group homogeneity among the characterization of several modules, which differentiates clearly in contrast to the other patterns. These differences can be explained by the basic orientation of the crowdfunding service systems, the differentiating motivation, risk, and legal requirements.

The configuration framework can be applied for both, the design of new crowdfunding service systems and the analysis of existing ones. In order to apply the configuration framework, each module has been assessed according to the desired output of the service system. The dominant design patterns may serve as a starting point. The parameter characteristics are designed as variations. Therefore, one can chose multiple variations for each parameter.

This paper contributes to research on crowdfunding and service science and provides especially two theoretical contributions. First, the paper expands crowdfunding research by proposing a crowdfunding service configuration framework, which describes the building modules of a crowdfunding service system and three dominant design patterns. Thus, the configuration framework and the dominant design patterns provide an overview over required service modules and respective parameters. By empirically deriving the dominant design patterns, thus verifying the appearance of certain design modules in specific contexts, we provide insights in the differentiating designs of crowdfunding service systems. This indicates that specific contexts (altruistic, hedonistic, and profit-oriented) require different modules in order to perform the context-specific service provision. By providing empirical evidence this paper supports and extends previous purely conceptual research on the modular structure of crowdfunding [6, 26, 27]. Further, considering the variety of crowdfunding service systems, the configuration framework may allow for the comparison of crowdfunding service systems on both, a functional and a component perspective, which might provide interesting results for a better understanding of crowdfunding in general and the design of crowdfunding service systems. Besides, the crowdfunding configuration framework possess predictive quality as the dominant design patterns indicate both, intra-group homogeneity and inter-group heterogeneity. Thus, the design patterns can be applied in order to predict the classification of a crowdfunding service systems to a certain crowdfunding archetype.

Second, the paper bridges research on crowdfunding with the field of service science. Thus, we contribute to service science, by proposing a framework for the systematic design of modularized services systems, which has been instantiated on the example of crowdfunding. Therefore, we contribute to the call for the design of novel artefacts, facilitating the engineering and management of service systems [4].

For practice, this paper provides guidance for potential providers to systematically design crowdfunding service systems. Further, it enables the decision support for the assessment of required competences, the identification of white spots for business opportunities, and a better understanding of the disruptive potential of crowdfunding. The three dominant design patterns serves as an initial blueprint for the implementation of a crowdfunding service system. Besides encouraging new market entrants e.g., banks or start-ups to systematically exploit white spots of the crowdfunding market and to develop new crowdfunding offerings, our findings might support established providers of crowdfunding service systems to evaluate their current system configurations.

We hope our study will encourage future research to take up the idea of crowdfunding as modular service systems. This might facilitate future studies to analyze the building modules of these service systems and their interrelations in more detail.

Table 3 Crowdfunding Service Configuration Framework

Service Modules	*Parameters*	*Variations*			
Market Differentiation	Crowd Motivation	Altruism	Hedonism	Profit-Orientation	
	Specialization	Sustainability & Social Action	Startup & New Business	Private Consumption	Creative Projects & Products
	Compensation	Greater Good	Reward	Interest	Profit Share
Matchmaking	Capital Giver	Individuals	Institutional Investors		
	Capital Seeker	Individuals	Non-Profit Organizations	Non-Governmental Organizations	For Profit Organizations
Crowd Activation	Offline	None	Mass Advertising	Personalized Advertising	
	Online	None	Mass Advertising	Personalized Advertising	
Customer Support	Capital Giver Support	None	Offline Support	Online Support	
		Personalized Support	Automatized Support	Peer-to-Peer Support	
	Capital Seeker Support	None	Offline Support	Online Support	
		Personalized Support	Automatized Support	Peer-to-Peer Support	
Investor Relations	Communication Channels between capital givers/seekers	None	Traditional Communication Channel (E-Mail, Telephone, Fax etc.)	Modern Communication Channels (Social Media, Blog)	
	Performance Monitoring	None	Progress Bar	Portfolio Management System	
Contracting	Terms and Conditions	None	Standardized Terms of Use	Privacy Policy Regulations	Payment Regulations
	Legal Relationships after Funding Success	Directly between Capital Seekers and Givers	Indirect (via financial intermediaries e.g., banks)		
Risk Assessment	Due Diligence	None	Traditional (personal data & documents)	Data Analysis	
	Feasibility	None	Business Plan / Project Plan	Prototype	
IT Functionality & Operations	Platform Development & Hosting	In-House	External Service Provider	White-Label Solution	
	Registration Process	None	Website Login (E-mail & Password)	Social Login (Facebook/Google)	
	Applications	Web Application	Mobile Application		
Payment	Forms of Payment	Offline Payment	Traditional Direct Payment	Online Direct Payment	Direct Debiting
	Time of Payment	Pre-paid	Instant-paid	Post-paid	
	Debt Default Actions	None	Notifications	Dunning	Debt Collection
	Payment Processing	Directly between Peers (capital seeker and giver)	Indirect via Financial Intermediaries		
Authentication	KYC Capital Giver	None	Personal Offline Identification	Automated Digital Identification	Personal Online Identification
	KYC Capital Seeker	None	Personal Offline Identification	Automated Digital Identification	Personal Online Identification

References

1. Welfens, P.J.J.: Finanzinnovationen, Wachstum und transatlantische Bankenkrise. In: Baumann, W., Braukmann, U., Matthes, W. (eds.) Innovation und Internationalisierung, pp. 303-326. Gabler (2010)
2. Blohm, I., Sieber, E., Schulz, M., Haas, P., Leimeister, J.M., Wenzlaff, K., Gebert, M.: Crowdfunding 2020 - Komplement oder Substitut für die Finanzindustrie? BoD – Books on Demand, Norderstedt (2015)
3. Michels, R., Hoffmann, V.: Crowdfinanzierung in Deutschland. vol. 4/2016. Für-Gründer.de (2016)
4. Böhmann, T., Leimeister, J., Möslein, K.: Service Systems Engineering. Business & Information Systems Engineering 56, pp. 83-90 (2014)
5. Böhmann, T., Krcmar, H.: Modulare Servicearchitekturen. In: Bullinger, H.-J., Scheer, A.-W. (eds.) Service Engineering: Entwicklung und Gestaltung innovativer Dienstleistungen, pp. 377-401. Springer Berlin Heidelberg, Berlin, Heidelberg (2006)
6. Haas, P., Blohm, I., Peters, C., Leimeister, J.M.: Modularization of Crowdfunding Services – Designing Disruptive Innovations in the Banking Industry. 36th. International Conference on Information Systems (ICIS), Fort Worth, USA (2015)
7. Hevner, A.R.: A three cycle view of design science research. Scandinavian journal of information systems 19, pp. 4 (2007)
8. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28, pp. 75-105 (2004)
9. Peters, C., Kromat, T., Leimeister, J.M.: Complex Services and According Business Models – Design and Evaluation of an Analysis Framework in the Field of Telemedicine. 48th Hawaii International Conference on System Sciences (HICSS) Koloa, Hawaii, USA (2015)
10. Chesbrough, H., Spohrer, J.: A research manifesto for services science. Communications of the ACM 49, pp. 35-40 (2006)
11. Maglio, P.P., Srinivasan, S., Kreulen, J.T., Spohrer, J.: Service systems, service scientists, SSME, and innovation. Communications of the ACM 49, pp. 81-85 (2006)
12. Menschner, P., Peters, C., Leimeister, J.M.: Engineering knowledge-intense, person-oriented services-A state of the art analysis. pp. (2011)
13. Edvardsson, B., Skålén, P., Tronvoll, B.: Service systems as a foundation for resource integration and value co-creation. Review of Marketing Research 9, pp. 79-126 (2012)
14. Böhmann, T., Krcmar, H.: Modulare Servicearchitekturen. In: Bullinger, H.-J., Scheer, A.-W. (eds.) Service Engineering, pp. 377-401. Springer, Berlin, Heidelberg, Germany (2006)
15. Balzert, H.: Lehrbuch der Software-Technik. Spektrum, Akadamischer Verlag, Heidelberg, Berlin, Oxford (1996)
16. Peters, C., Leimeister, J.M.: TM3-A Modularization Method for Telemedical Services: Design and Evaluation. In: Proceedings of 21st European Conference on Information Systems (ECIS). Utrecht, Netherlands (2013)
17. Tuunanen, T., Cassab, H.: Service Process Modularization. Journal of Service Research 14, pp. 340-354 (2011)
18. Voss, C.A., Hsuan, J.: Service Architecture and Modularity. Decision Sciences 40, pp. 541-569 (2009)
19. Blohm, I., Leimeister, J.M., Krcmar, H.: Crowdsourcing: How to Benefit from (Too) Many Great Ideas. MIS Quarterly Executive 12, pp. (2013)

20. Belleflamme, P., Lambert, T., Schwienbacher, A.: Crowdfunding: Tapping the right crowd. Journal of Business Venturing, Vol. 29, No. 5, pp. 585-609 (2014)
21. Bradford, C.S.: Crowdfunding and the Federal Securities Laws. Columbia Business Law Review, Vol.2012, No. 1 (2012)
22. Haas, P., Blohm, I., Leimeister, J.M.: An Empirical Taxonomy of Crowdfunding Intermediaries. In: International Conference on Information Systems (ICIS). Auckland, New Zealand (2014)
23. http://www.crowdsourcing.org/editorial/2013cf-the-crowdfunding-industry-report/25107
24. Tomczak, A., Brem, A.: A Conceptualized Investment Model of Crowdfunding. Venture Capital 15, pp. 335-359 (2013)
25. Wieck, E., Bretschneider, U., Leimeister, J.M.: Funding from the crowd: An internet-based crowdfunding platform to support business set-ups from universities. International Journal of Cooperative Information Systems (IJCIS), Vol. 22, No. 3, pp. 1-12 (2013)
26. Liebenau, J., Elaluf-Calderwood, S., Bonina, C.: Modularity and network integration: Emergent business models in banking. In: 47th Hawaii International Conference on System Sciences, pp. 1183-1192. (2014)
27. Hemer, J.: A Snapshot on Crowdfunding. Working papers: Firms and Region (2011)
28. O'Sullivan, J., Edmond, D., Ter Hofstede, A.: What's in a Service? Distributed and Parallel Databases 12, pp. 117-133 (2002)
29. Zwicky, F.: The morphological approach to discovery, invention, research and construction. New methods of thought and procedure, pp. 273-297. Springer (1967)
30. Zwicky, F., Wilson, A.G.: New methods of thought and procedure: Contributions to the symposium on methodologies. Springer Science & Business Media (2012)
31. Hartmann, P.M., Zaki, M., Feldmann, N., Neely, A.: Big data for big business? A Taxonomy of Data-driven Business Models Used by Start-up Firms (2014)
32. Peters, C., Blohm, I., Leimeister, J.M.: Anatomy of Successful Business Models for Complex Services: Insights from the Telemedicine Field. Journal of Management Information Systems, Vol. 32, pp. 75-104 (2015)
33. Peters, C.: Together They are Strong - The Quest for Service Modularization Parameters. In: European Conference on Information Systems (ECIS), Tel Aviv, Israel (2014)
34. Österle, H.: Business Engineering. Prozess- und Systementwicklung: Band 1: Entwurfstechniken. Springer-Verlag (2013)
35. Gordijn, J.: E³-value in a Nutshell. In: International Workshop on E-Business Modeling. (2012)

Service Modularization in an Evolving Context: A Comparison between the Old and the New World of Automotive Engineering Services in Germany

Florian Müller

University of Bremen, Faculty of Business Studies & Economics, Bremen, Germany
mueller@is.uni-bremen.de

Abstract. Modularity in services has emerged as a promising design approach that offers a sound balance between cost-efficiency in service production and customization. However, the existing literature draws on a narrow scope of service design that considers certain characteristics of service as constraints to the successful adoption of the concept, i.e. service modularization. The growing importance of service as a central logic to explain value creation suggests the need to expand this view and to understand how service modularization and changes in dynamic service contexts may come together. This article makes a first attempt in this direction by exploring service modularization in the light of major restructurings in the field of automotive engineering services (AES) in Germany. Through 22 qualitative interviews with customers and providers of AES, this article presents a conceptual process model that explains how service modularization of AES is propelled by transformational forces in an evolving service context.

Keywords: Service Modularization, Service Modularity, Automotive Engineering Services, Evolving Context, Service Logic

1 Introduction

The increasing significance of service in our modern economy [1] is accompanied by a growing pressure on service providers to become more cost-efficient in service provision. On the other hand, customers are more demanding than ever and expect customized services for their individual needs [2]. Consequently, ambitions to achieve cost reductions and efficiency gains through service standardization [3] are contradicted by the fact that providers need to remain flexible to meet the heterogeneous demands of different customers at once.

From this challenge, service modularity has emerged as a promising design approach to achieve a sound balance between standardization and individualization at the same time [1]. In general, modularity can be viewed as an architectural principle that describes the decomposition of a complex system into smaller parts, i.e., modules that are characterized by a high internal cohesion but are loosely coupled among each other [4]. Standard interfaces between modules enable a high flexibility in the system and ensure that modules can function together as a whole [4].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Müller, F. (2017): Service Modularization in an Evolving Context: A Comparison between the Old and the New World of Automotive Engineering Services in Germany, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 166-180

As of today, services are rarely considered to adhere to a modular design (except for specific industries like, e.g., logistics and financial services) [1] and the question of the extent to which different services are eligible for an adoption of the concept of modularity, i.e., modularization, is subject of a controversial discussion in the literature [5, 6]. Contrasting viewpoints have especially emerged from an evolving understanding of service. Currently, a large part of the literature limits the role of modular design to the level of service operations. From this narrow scope, certain service characteristics are considered constraints to modularization, e.g., an active customer in service processes [7, 8]. However, the current shift of the academic community towards a service-centric view of the world [5] suggests the need to expand the current view on service modularity and to elaborate on our understanding of how service modularization may take place in the dynamic service contexts of the real world [9, 10].

This article makes a first step towards this direction and explores interrelations between modular service design and an evolving service context at the case of automotive engineering services (AES) in Germany. AES are offered in a project-based fashion by professional service firms to large original equipment manufacturers (OEM) of automobiles for the design, production and testing of technical models for vehicle parts. Recently, industry studies outline a growing dynamic in the automobile industry and changes in the design of AES projects [11]. Considering the need for AES providers to balance contrasting aims of customization and standardization, modularity could be a solution. However, little is known on how contextual changes lead to a higher degree of modularity in services, thus limiting opportunities for modularization in dynamic sectors.

Consequently, the key objective of this research is to identify factors in the broader service context in which AES are developed and rendered that may facilitate modular AES design. Furthermore, the article aims to develop a conceptual model that explains such interrelations. As encouraged by past research, particular attention is devoted to the business relation between customers and providers [1, 12]. To tackle our research aim, we employ a qualitative field study among customers and providers of AES.

The paper begins with a brief outline of our theoretical understanding of service modularization and a brief description of the abductive research methodology. Subsequently, we present an analysis of transformational forces that cause changes in different interconnected layers and facilitate the modularity of AES. Chapter 5 explains how these changes in the design of AES come together with an elevation of the business relation between providers and OEMs. Our findings are summarized in a conceptual process model that illustrates causal relations between different layers of change, before we conclude with a discussion of research implications and future research avenues.

2 Research Background

The paradigm shift towards a service-centric view of service as a central logic to explain value creation raises tensions with the traditional assumption that certain service characteristics constrain modularization. Tensions are especially evident with respect to Service Systems Engineering (SSE) that adopts a systemic perspective at service as a socio-technical system that enables actors and resources to co-create value [13, 14].

The interaction of different parties with a service system allows for a systemic design. However, design choices in the practice are often determined by complex and rapidly evolving business environments [9], so that modularization may create tensions [7].

With respect to modular service design, the emergence of modular financial services [15] has shown that modularization processes may be triggered by changes in the service context, e.g. through technological advances or changes in legal regulations. Yet, little is known on how contextual changes propel modularization processes in the practice. Research on the interdependencies between modularization and evolving contexts seems especially important since scholars are now increasingly concerned with the design of services that are more adaptive to changes in dynamic contexts.

Service modularization is considered the decomposition of a service into smaller parts and a restructuring of these parts in form of service modules [2]. However, the nature of service is highly complex and modular service design shows in various forms [16] and in different dimensions [5]. This study draws on a well-established understanding of modular service design from Pekkarinen & Ulkuniemi [17] who suggest to examine service modularity in three different dimensions (*3D of modularity*): *Modularity in the service organization* refers to the organization, integration and use of internal or external resources in the service production. *Modular service processes* are built from modules based on information processes or physical operations. *Modular service offerings* relates to service elements whose value is separable and visible to the customer.

The industry context in which AES projects are designed and rendered is currently subject to an increasing dynamic. Providers of AES face growing competition and cost-pressure that lead to consolidations in the market [18]. Nowadays, the remaining players in the market offer a wide array of engineering services to a low number of OEMs that traditionally expect customized solutions for their individual needs. For the providers service modularization would be the concluding solution. However, little is known on modular service design in the AES sector. While in automobile manufacturing and supply chain processes, modularity has been a popular concept for many decades [1, 19], prior research outlines several characteristics of the AES sector as natural constraints to modular service designs, e.g., the reliance on complex problem solving during service provision and a close interaction between customers and providers [7, 8]. However, recent industry studies outline the transformative role of digitalization in the AES sector and indicate ongoing evolution in the design of AES projects that could lead to a higher degree of modularity in services [11, 18]. The dynamics of the AES sector create a particular interesting setting for an exploration of contextual variables that influence the application space for modular service design.

3 Methodology

In order to identify facilitators of service modularity in an evolving business environment – a phenomenon that is highly complex and insufficiently understood – a qualitative research design was chosen due to the explorative nature of our research aim. After familiarizing with the current body of knowledge on modularity in services, the research followed an abductive methodology which has been emphasized to be particularly useful to study ill-structured phenomena and to discover new things [20].

Consistent with an abductive approach to case research, collection of data, review of literature and the development of a framework for the analysis of the data were carried out in an intertwined fashion as part of an iterative process. Dubois and Gadde [20] describe this process as systematic combining and matching of existing theory, research framework, empirical world and findings. During this process the researcher moves back and forth between field observations and theoretical concepts in order to enhance both understanding of the field of research and the empirical data [20].

Data collection took place in one phase between December 2014 and June 2015. During this period, 22 semi-structured interviews were conducted among high-level executives within the German automobile industry. Potential interviewees were identified through an industry-specific key word search in the database of a public business network for employee roles and positions that were involved in the rendering of AES projects, which represents a form of judgement sampling [21]. Interviewees were invited to the research by email. This approach resulted in a response rate of 21 percent.

Interviews were carried out personally or by telephone and lasted between 30 and 60 minutes. From the side of AES providers interviewees were mostly CEOs and Business Unit Managers. From the side of the manufacturers interviewees were working as engineering, process or purchasing executives. All interviewees possessed deep industry knowledge due to a long work history (more than 10 years) in the automobile industry. Table 1 gives an overview of the sample characteristics. To increase the reliability of the data, interviews were audio-recorded and transcribed in German language [22] before later during the writing process, relevant quotes were translated into English. Systematics in the analysis and interpretation of the interview transcripts was supported by employing MaxQDA, a professional software for qualitative data analysis that was used to assign codes to relevant text passages in the interview transcripts.

Table 1. Sample Characteristics

Perspective	*Engineering service provider (P)*							*OEM (C)*			
Employees	0-1000		1001-5000			5001-15000		>15001			
Firm	P1	P2	P3	P4	P5	P6	P7	C1	C2	C3	*Total*
# Interviews	1	3	2	2	2	1	1	5	3	2	22

As given by the abductive approach, collection, analysis and interpretation of empirical data were carried out in parallel to the development of the scope and aim of the research and the research framework, so that the interview guidelines and the coding scheme evolved in the course of our study through multiple iterations of application, interpretation and adaption [20].

At the beginning of this process, a tentative research framework was derived from the literature on service modularity that comprised basic interview questions and a tentative coding scheme for the systematic analysis of the data. The first set of interview questions aimed to improve our general understanding of the design of AES and the dynamic context in which AES are offered and produced. Complementary perspectives of customers and providers at AES were considered to support this aim. Correspondingly the tentative coding scheme allowed us to examine the adherence of AES to basic principles and effects of service modularity e.g., loose coupling, reuse, bundling, flexibility and standardization [16, 23], in each of the 3 dimensions of service modularity

[17]. This approach gave our interpretation of the data orientation and still allowed the necessary flexibility to explore modularity in AES with an open mind.

During the research process the scope of the research and the research framework were narrowed and adopted to our findings. To give an example, when we denoted that several interviewees highlighted causal relations between modularization and regulatory changes, the coding scheme was enhanced with new codes to examine this relation throughout all interview transcripts. Likewise, over time codes were added to classify the variety of changes in the context of the AES sector into different layers (industry context, service design and business relation). Layers are an outcome of the abductive approach and help to lower the complexity in the outlined industry transformation. Further codes were added to capture causal relations between layers and temporal reference of statements (past, present and future). A continues comparing of findings with the literature on service modularity revealed a research gap with respect to how modularization processes interrelate with changes in evolving contexts.

Finally, after the aim of the research and the framework were refined in the course of a continuous direction and redirection of the study, the research process was ended when the scope of the research, the applied research framework and the empirical findings gave a sound and coordinated impression. At this point, interviews yielded only little additional insights and indicated a sufficient theoretical saturation [20]. In a last step, the findings were summarized and integrated into a conceptual process model.

4 Transformational Forces and Facilitators of Service Modularity

Interviewees from providers of AES and OEMs outline an ongoing transformation of the global automobile industry that manifests in a rapidly evolving business environment of the AES sector. Consistent with prior studies [11, 18], it has become evident that processes of digitalization and socio-economic change are major forces behind the current restructuration. In the course of our research, we denoted that changes in the AES industry could be located on three different interconnected layers: layer of the *industry* (1), *service design layer* (2) and, layer of the *business relation* (3). The layers evolved in the course of the abductive research approach. This chapter provides a brief outline of the three different layers of change and subsequently with a more differentiated understanding of key transformational forces.

The first layer comprises changes in the context of the automobile industry (industry layer). Changes concern multiple actors in the market and e.g., lead to AES providers becoming less dependent on the OEMs in terms of how they render their services.

Second, several changes can be denoted in respect to the design and rendering of AES projects (design layer) that allow for standardization, reuse and decoupling of service elements that were closely intertwined and tied to the approval of the OEMs in the past. Such changes were considered indications of an ongoing modularization of AES.

Third, we observe a shift in different socio-psychological characteristics of the *business relation* between AES providers and OEMs that elevates the longitudinal exchange relation between the parties onto a new level, e.g., in terms of trust and control.

Interviewees suggest a close relation between changes in different layers that connect to the increasing adherence of AES to principles of modularity. To gain deeper insights into how changes relate across different layers and to identify changes in the context that facilitate modularization processes in the practice, next we present a more differentiated view at four major transformational forces that were outlined by the interviewees in close relation to changes in the design of AES:

- *OEMs adapt their traditional core competencies (development and manufacturing)* to new challenges of the digital age. This is especially shown by the OEMs increasing their investments into IT-related competencies, i.e., software development, while they leave the amount of technical engineering capacity on a constant level. Since the number of car variants has spiked in recent years and leads to an ever-growing demand for engineering capacity in the market, OEMs outsource more and more of their engineering activities and allow their providers to play roles that are more important in the automotive value creation chain.
- *Engineering projects grow in complexity and extent.* In the past, the OEMs' outsourcing strategies aimed at creating a large number of smaller engineering projects in order to yield cost-reductions. Today, OEMs attempt to bundle closely related engineering activities to larger project extents. Interviewees attributed shift in the outsourcing strategy to a rise in the complexity of in-car electronics, i.e., sensors and IT in modern automobiles that demand for a closer alignment of different engineering activities and increases the effort for the management and integration of different engineering projects.
- *A change in the legal regulations of service contracts* demands for a clearer separation between real project contracts and body leasing. Despite a clear differentiation between the workforce of customer and provider is requested by the law, this request was loosely handled in the past. It has been a common practice for AES providers to deploy their engineers directly on site at the OEMs. Interviewees outlined that the control of the legal request for a clearer differentiation would now suddenly be taken serious. As a first effect, service contracts now need to be specified in more detail. As a second effect, providers have to lower the proximity to their customer by relocating a majority of their workforce into their own office spaces.
- In recent years, *AES providers have shown a significant growth* in turnover and in the number of their employees. This growth is paralleled by an *increasing diversification* in the customer portfolio of many providers and in the variety of services offered. Several providers outlined opportunities from them to yield synergies from cross-customer learning and a growing dependency of the OEMs from external engineering expertise in certain technological fields related to AES.

To underline the strength and velocity of these transformational forces, several interviewees differentiated between an *"old world"* and a *"new world"* of AES. In this respect, a major breaking point between both worlds can be located between the years 2014 and 2015. The following sections present an analysis of AES design in the light of the outlined industry transformation in three dimensions of modular service design. The terminology (old world/new world) is thereby adopted to compare two contrasting service settings that determine the potential for modular service design of AES.

4.1 Service Modularity in AES Processes

AES processes are traditionally characterized by a high degree of interaction between customer and provider who co-produce the service in close collaboration. During the analysis of our interview data, we observed changes in the interaction between the parties as well as in the controle and influence that the parties exert at the course of AES processes. The first change concerns *the authority over the extent and the intensity in the interaction*. In the old world, customers traditionally decided over their own role in the service process. In this role, they fulfilled a strong supervision function and determined the level of their involvement during service production. In the new world, providers and customers negotiate a work mode that lowers the controlling effort for the customer and gives the provider longer times of autonomous service production.

"Now, we have to find a balance between working in quiet without being disturbed by the provider and reporting to the customer as that we nowadays have to control the intervention of our customers. […] we had to establish our own reporting and project management processes to offer our customers more orientation, so that any developments for them do not come as a surprise. In this way, we impede that the customer interferes on a regular base, as it was a common practice in the past." (P1, CEO)

Interviewees stated that OEMs increasingly expect their providers to manage the necessary level of interaction between the two parties (e.g., for decision-making). In the new world, providers need to find an efficient level of interaction that is beneficial for both parties. At the same time, providers need to ensure a sense of control for their customers and update them with reliable information on the project progress on a regular base. Consequently, we see an enlargement of the interval of *customer-provider interaction*. In the old world, interviewees outlined a continuous dialog during service production and discussion on a low level of technical detail that led to many serial iterations through previously completed process steps and serial discussions. Providers expressed the need to concentrate the interaction at shorter and more efficient periods, i.e., to organize the interaction in a more *punctual* manner.

As a third change, OEMs handover some of the *service process authority* to their providers. In the old world, providers engineering activities were closely supervised by the OEM. Providers thus, could only exert a limited influence at the technical solutions. In particular, the alignment of dependencies between design decisions and other parties (e.g., other projects and departments) was usually under OEM authority. Interviewee comments suggest that customers in the new world are increasingly handing over this responsibility to their providers, who need to manage interdependencies, carry the risk over design decisions and thereby increasingly produce the service in full authority.

Table 2 summarizes the outlined changes in the process dimension. In the new world providers increasingly manage the interaction between them and their providers and gain more control over the customer interference with the actual engineering work. This encloses a larger space to providers for the modularization of internal processes for the development of technical drawings. In the new world, providers find more space for process standardization and have more freedom in the organization of (mostly human) process interfaces. Standardization is thereby introduced in a part of the value creation process that remains hidden from the customer, e.g., through automatization of repetitive tasks and the introduction of standards in the setup of internal projects. In addition, the interaction with the customer becomes more standardized. More process authority allows providers to introduce internal standards into controlling and reporting processes and the development of replicable routines for the preparation and presentation of alternative solutions to the customer. These service elements are decoupled from the customized parts of the service and thus become available for use in different service projects. By drawing on combinations of standardized and customized services processes, providers maintain the necessary flexibility to meet individual expectations of the customer with respect to how the interaction between both parties is organized, while also increasing efficiency in major parts of the value creation process.

Table 2. Facilitators for achieving modularity in AES processes

Facilitator	*Old world*	*New world (trend)*
Authority over customer-provider interaction	Customer	Provider
Interval of customer-provider interaction	Short	Longer (punctual)
Service process authority	Customer	Provider

4.2 Service Modularity in AES Organizations

AES providers typically compose their internal resources (the majority consisting of human engineering capacity) along two different organizational dimensions. The first dimension mirrors key phases of the automobile product development process. Providers draw on competence teams that are organizational units specialized in phase-specific requirements and demands, e.g., competence teams possess knowledge and skills of advanced development, series development or simulation and assurance. As a second dimension, providers align their competences with different technological spheres of the automobile, e.g., interieur, exterieur or chassis development. Consequently, the combination of both dimensions (process and technology) forms a matrix organization. In theory, this service organization allows combinations of competencies along the technological and process dimension. However, in practice providers outlined several limitations with respect to opportunities to offer services based on combinations of different competence fields. Interviewees outlined additional combinatorial constraints that hardened the integration and connection of separate competences in the old world. However, we see several of these obstacles being reduced or even eliminated in the new world.

The first change concerns the *physical proximity between customer and provider*. AES providers typically draw on a decentralized engineering capacity that distributes across different service branches from which most lie in close distance to the development and manufacturing centers of the large German OEMs. In the old world, providers' had to integrate physically into the organization of their customers to enable co-production of the service on site of the customer in close collaboration. Thus, an important premise for successful service provision was that a large part of the provider's engineering workforce had deep knowledge of the customers' internal processes and was highly familiar with the decision-making culture of the customer.

"Until now many services were offered as projects in pretense, as that in theory responsibility and authority over our employees lay with us, but in reality it laid with the customer. Our workforce was at the customer's site, they were sitting in the customer's office space and the customer directed them to do this and to do that. […] At the moment there is a big movement, so that the manufacturers seek to reach a spatial

separation between them and their external engineering providers." (P3, BU Manager)

The number of provider employees working on site of the customer has significantly lowered in response to a recent tightening of compliance regulations and providers gain more authority over the course and outcome of their services:

"It is increasingly our decision which work packages we produce and where we do this. This is a major change that one can currently observe in the market." (P7, CEO)

This reduction of the proximity between the workforce of providers and OEMs leads to a concentration of the interaction between both parties at fewer *(human) interfaces*. In the old world, OEMs often discussed technical and project management decisions directly on the lower levels of the provider hierarchy. In the new world, much of the communication runs through a considerable lower number of well-networked intermediaries. Project managers often fulfill this function and manage the communication between their own project teams and the OEM. They often maintain a personal relation to relevant stakeholders that allows them to facilitate decision making within the internal organization of their customers. Finally, interviewees denote a "professionalization" of the IT and data management systems (DMS) for the organization and exchange of technical models and project related data but also a growing heterogeneity of these systems. While in the old world, customers usually had the authority over the DMS and providers were connected through remote- or on-site-access, OEMs in the new world often expect their providers to have their own IT-infrastructure in place and to manage the data exchange in full authority. Other interviewees expect a growing importance of third party DMS-solutions in the future. In this respect, IT-security, efficiency in the collaboration and compliance regulations are seen as contrary aims that yet remain an unsolved problem. To become more flexible and less dependent from OEM-specific premises providers are beginning to invest into their own IT.

The outlined changes (cf. Table 3) lower the barriers for providers for drawing upon their full range of competencies and engineering capacity when offering services to different OEMs. In the new world, it is increasingly the decision of the provider where and by whom certain engineering activities are carried out, so that value-creating activities are decoupled from the location where services are offered. This suggests the emergence of modular organizations that enable providers to offer a wider set of competencies at different service branches irrespectively of where the underlying engineering capacity are located. Providers can build upon their workforce from different corporate sites to put up the necessary capacities and the variety of competencies that are required to handle the complexity of larger projects in full authority. Modularity in the organization is also shown by the fact that providers stated to outsource a growing amount of repetitive engineering activities to other countries to yield cost reductions - a trend that is strengthened further by lowering barriers for the exchange of data.

Table 3. Facilitators of modularity in AES organization

Facilitator	*Old world*	*New world (trend)*
Physical proximity	High (integrated)	Low (dispersed)
Extent of customer-provider interfaces	Many interfaces on a low level-hierarchy	A few higher-level intermediaries
Data management authority	Centralized at the customer	Decentralized

4.3 Service Modularity in AES Offerings

AES are typically offered as highly customized service projects. Customization proceeds throughout the course of the service process and begins with a tender and proposal phase. Interviewees characterized the process of customization as a reciprocal dialog between customer and provider in which both parties negotiate the value of the service offering. Our findings show three significant differences between the old and the new world with respect to the process of customization and the decomposability of AES offerings.

First, several interviewees outlined that in the old world providers were mainly offering their engineering capacity to the customer and thus had only limited opportunities to offer combinations of their competencies. Providers indicated that the growing complexity and extent in AES projects further propel a shift in the focus of their offerings from capacity towards competencies:

"In simple words, we observe a transition from a focus at [engineering] capacity to competence. […] When projects become larger, the customer's confidence into the provider becomes more important, particularly into the providers' project management capabilities" (P3, Business Unit Manager)

Consequently, providers see growing opportunities to offer combinations of different competencies and to bundle engineering activities to larger service packages, which in the old world were offered as separate projects. Interviewees highlighted that their customers would increasingly expect them to enrich the value in their offerings through the integration of additional support services, such as more sophisticated project management processes that can lower the customers' controlling effort.

We further observe a change in the influence that providers exert at the design and the extent of service offerings. Analysis of our data revealed that in the old world, the boundary conditions for later service provision (e.g., interfaces to other projects and project content) were merely determined by the customer independently of the provider. Providers were mainly concerned to design an initial offering (delivered in written form as a project proposal) that matches the customers' requirements. In contrast, several interviewees outlined that customers would increasingly invite providers to integrate themselves during the specification of the tender, allowing them to participate in the fundamental design of their service offerings in a very early phase:

"What now defines our work is early integration of suppliers. […] Many suppliers offer this preparatory effort for free to take predominant positions during the commissioning phase. Then the coming tasks are put to the agenda and we consider in common who performs which task by oneself or in cooperation with a partner." (C2, Chief Engineer)

Finally, the customization process of service offerings develops from a persistent refinement towards a guided implementation of dynamically created configuration options during which the offering evolves incrementally from different customer choices. Providers and customers switch between phases of requirement specification, preparation of alternative options and subsequent decision making with the customer, which is in many ways a different work mode compared to the old world, where the solution went through a less structured process of direct and continues refinement.

The three outlined changes (cf. Table 4) facilitate modularity in AES offerings.

Providers stated that the offering of combinations of their competencies becomes a key to remain competitive in the marketplace. In this regard, providers outlined the importance for them to display the value inherited in different combinations as part of larger offerings. Providers gain new opportunities to influence project requests before official commissioning and can thereby incorporate interfaces between project parts that in the old world were often distributed among different providers. In this way, providers described ways to implement internal standards for collaboration and communication to increase efficiency in service production. On the other hand, providers point out the need to obscure the true degree of modularity in their service offerings to their customers in order to reap internal synergies. Competencies in the AES industry can be viewed as service modules, whose evaluation, selection, integration and rendering is part of a collaborative customization process between customer and provider. In the new world, this process is often guided by the provider.

Table 4. Facilitators of modularity in AES offerings

Facilitators	*Old world*	*New world (trend)*
Role of provider competencies in the offering	Lower	Higher
Provider influence on service project design	Lower	Higher
Nature of the customization process	Continuous	Incremental

5 Regulating Variables of Service Modularity

Changes in the design of AES – such as a lowering of the physical proximity between providers and OEMs – come together with a shift in the business relation between the two parties. In this respect, interviewees outlined changes in various socio-psychological factors, e.g., in terms of trust, expectations, appreciation, control and responsibility. While in the old world, providers were merely regarded as enhancements of the engineering capacity of the large OEMs, interview comments such as the following suggest that the relationship between the two parties elevates onto a higher level and the emergence of genuine engineering partnerships:

“The relation between engineering service providers and the manufacturers changes from one important point of view. One cannot speak about normal suppliers anymore, instead it must arise partnerships at eye level. […] Providers have to prove themselves for real partnerships to emerge.” (C1, Head of Engineering Processes)

OEMs are now narrowing the scope of their own technical development activities and thus need to share larger proportions of project-related authority and risk to handle the complexity of engineering projects. This makes OEMs more dependent on the skills and resources of their providers. Interviewees stated that OEMs have to let loose when commissioned project amounts are increasing and that they need to have more faith into their provider’s competencies and capabilities to carry out projects to full extent in own responsibility. Deeper analysis of our interview data suggests four major changes on the layer of the business relation.

First, providers outlined that their customers would increasingly expect them to act as advisors who demonstrate a yet unknown level of *sovereignty* in their services provided. This achieve this, providers need to develop a more profound understanding of interdependencies between their individual engineering projects and the complex project environment of the OEMs in which these projects are embedded.

The second change concerns the level of *confidence* of the customer into the capabilities of the provider. A high manager of an OEM indicated that an elevated level of confidence would be a premise for them to lower their control over their providers' engineering activities during service provision. Providers who seek to take over larger and more complex extents have to prove themselves as reliable and capable to manage the complexity of interdependencies within the project environment of their customers in own authority and create high service value over longer periods.

Thirdly, in return customers stated that they would demand from their providers a stronger *commitment* to take over larger financial risks in engineering service projects and manage these risks in own authority. Providers confirmed this and indicated an increased need for the development of improved risk management techniques to handle risks in service projects but also on the level of the project portfolio.

Finally, we see an increase in the *recognition* of the providers' competencies in specific technical fields of knowledge, e.g., in the field of lightweight construction. OEMs are recognizing the value that lies in the value of competencies, instead of viewing them as sole engineering capacity.

Providers, who successfully manage change in these four regulating variables will encounter a larger freedom to design their service production according to their own account and based on principles of modularity. However, OEMs indicate that providers who fail to manage the transition towards genuine partnerships are endangered to become subcontractors by the larger providers. They cannot expect to reap the same potential of service modularity as those providers in the first-tier.

6 Discussion and Conclusions

This paper examines service modularization in the rapidly evolving AES industry and presents an empirically grounded comparison between two contrasting service settings, in which AES projects are developed and rendered. Findings are aggregated into a conceptual process model (Figure 1) that locates changes in the AES sector onto three different layers of change. The industry-specific process model outlines four key transformational forces in the context of the German automobile industry (1) that lead to the formation of modular service designs in the AES sector alongside each of the three modularity dimensions (2). However, modularization comes together with changes in the business relation between customers and providers of AES. Closer analysis shows that changes in the design of AES and changes in four different characteristics of the business relation are mutually dependent, so that these factors are considered to regulate the potential for modular service design (Layer 3).

As to the consequences of the transformation, controversial views can be denoted. Providers outline both positive as well as negative effects on them. On the one hand, they appreciate the development of genuine partnerships and opportunities for them to implement projects in full authority. Providers expect increasing opportunities to offer their services to new customers, e.g., to OEMs located in the US and in Asia. On the other hand, providers indicate that modularization enhances comparability of service offerings, which further reinforces the cost-pressure on providers that face increasing international competition. In this respect, some providers stated that they would mainly aim to achieve standards in service processes and organization but mask the full extent of modularity in their service offerings to reduce comparability in the marketplace. In general, modularity of AES is less a suitable strategy for providers to outperform competitors and rather becomes a prerequisite to remain competitive in the marketplace.

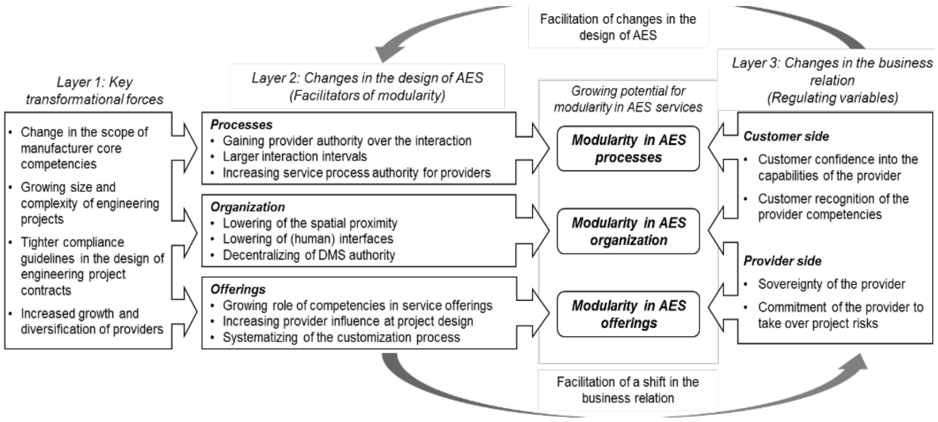


Figure 1. Causal relations between different layers of change and service modularity in AES

6.1 Theoretical and Practical Implications

Theoretical implications concern our understanding on the role of evolving contexts in service modularization processes. While existing research on service modularity is traditionally concerned with modularity in a specific context, this article is one of the first that examines how service modularization may be propelled by transformational forces in an evolving industry context [5]. The presented process model (Figure 1) offers an overview on several variables in the context of the AES industry that determine the potential for service modularity. Considering rapidly evolving contexts in our digital age [9], our findings underline the need to further elaborate on the role of modularity from a service-centric view [13, 14]. In this respect, in line with recent research in the field of service innovation, our findings underline the importance of modularity as an important theoretical design construct in the field of SSE [13, 14] that may contribute to our understanding of the relations and interactions of multiple actors in service ecosystems [10]. Contrary to the frequent assumption that modular service designs are necessarily an outcome of a systematic modularization initiative in a specific service context, the case of AES shows that an adoption of the concept, i.e. modularization may be driven by design choices created from changes in dynamic business environments. In this respect, networked business relations should receive further research attention to develop more comprehensive modular design strategies for services [1]. The presented facilitators of modularity on different layers of change offer an interesting launch point for such research endeavors.

As for the *managerial implications*, the presented conceptualization of three different interconnected layers of change outlines a path for AES providers to develop a modular service design while also becoming more independent of their customers. In order to manage the transition to the new world of AES successfully, providers have to facilitate and manage changes on the layer of their business relations and in the design of their services in close coordination. Providers in this relation highlighted the role of a wide-reaching network to decision makers within their customers' organization and improved communication skills on the level of project managers. The outlined differences between the old and the new world of AES can be regarded in this respect as a development target for the design of AES at which providers can align the facilitation of design changes along the three modularity dimensions and elaborate on the systematic reshaping of their business relations. However, findings also show that providers need to consider the extent to which they reveal the degree of modularity to their customers. A modular service design, whose architecture is fully transparent to the customer, may expose the provider to a high cost-pressure, because customers can easily compare service modules with other offerings in the marketplace.

6.2 Limitations and Future Research Avenues

Our study is beset with some *limitations* that can motivate future research. First, this study builds upon a large set of data that allowed us to identify general changes in AES. However, generalizability of our findings is likely to be limited to the German AES industry until additional empirical studies have validate the presented causal relations in other service contexts. Second, despite the concept of the three dimensions of modularity is widely recognized, it leaves room for interpretation in terms of what modules are comprised of and how they can be identified in the practice. Because a universal definition of service modularity has not yet emerged [5, 6], we differentiated between modularity in three dimensions (3D of modularity) [17] and enhanced this understanding by the general effects and principles of service modularity - in particular those defined by Dörbecker & Böhmann [16], Bask et al. [6] and Tuunanen et al. [23]. While, this broader interpretation of modularity in services allowed the identification of several modularity effects and principles during the analysis of the interview data, our findings suggest the existence of different forms of modular service design.

To summarize, this study contributes an empirically grounded conceptualization on service modularization in the e context of the German AES sector that explains how service modularization and changes in an evolving service context may come together.

References

1. Bask, A., Lipponen, M., Rajahonka, M., Tinnilä, M.: Framework for modularity and customization: service perspective. J. Bus. Ind. Mark. 26, 306–319 (2011).

2. Voss, C.A., Hsuan, J.: Service Architecture and Modularity*. Decis. Sci. 40, 541–569 (2009).
3. Böttcher, M., Klingner, S.: Providing a method for composing modular B2B services. J. Bus. Ind. Mark. 26, 320–331 (2011).
4. Baldwin, C.Y., Clark, K.B.: Design Rules: The power of modularity. MIT Press, Cambridge (2000).
5. Müller, F., Lubarski, A.: School of thought in service modularity. In: Proceedings of the European Conference on Information Systems (ECIS). , Istanbul, Turkey (2016).
6. Bask, A., Lipponen, M., Rajahonka, M., Tinnilä, M.: The concept of modularity: diffusion from manufacturing to service production. J. Manuf. Technol. Manag. 21, 355–375 (2010).
7. Cabigiosu, A., Campagnolo, D., Furlan, A., Costa, G.: Modularity in KIBS: The Case of Third-Party Logistics Service Providers. Ind. Innov. 22, 126–146 (2015).
8. Carlborg, P., Kindström, D.: Service process modularization and modular strategies. J. Bus. Ind. Mark. 29, 313–323 (2014).
9. Ostrom, A.L., Parasuraman, A., Bowen, D.E., Patricio, L., Voss, C.A.: Service Research Priorities in a Rapidly Changing Context. J. Serv. Res. 18, 127–159 (2015).
10. Chandler, J.D., Lusch, R.F.: Service Systems A Broadened Framework and Research Agenda on Value Propositions, Engagement, and Service Experience. J. Serv. Res. 18, 6–22 (2015).
11. Kleinhans, C., Neidl, T., Radics, A.: Automotive Entwicklungsdienstleistung. Verband der Automobilindustrie e.V. (VDA), Berlin (2015).
12. Rahikka, E., Ulkuniemi, P., Pekkarinen, S.: Developing the value perception of the business customer through service modularity. J. Bus. Ind. Mark. 26, 357–367 (2011).
13. Böhmann, T., Leimeister, J.M., Möslein, K.: Service Systems Engineering: A Field for Future Information Systems Research. Bus. Inf. Syst. Eng. 6, 73–79 (2014).
14. Maglio, P.P., Vargo, S.L., Caswell, N., Spohrer, J.: The service system is the basic abstraction of service science. Inf. Syst. E-Bus. Manag. 7, 395–406 (2009).
15. Allchim, C., Matt, A., Ted, M., Fine, A.: Modular Financial Services - The new shape of the industry. Oliver Wyman (2016).
16. Dörbecker, R., Böhmann, T.: The Concept and Effects of Service Modularity – A Literature Review. In: 2013 46th Hawaii International Conference on System Sciences (HICSS). pp. 1357–1366 (2013).
17. Pekkarinen, S., Ulkuniemi, P.: Modularity in developing business services by platform approach. Int. J. Logist. Manag. 19, 84–103 (2008).
18. Blöcker, A.: Branchenentwicklung Entwicklungsdienstleister. Hans-Böckler-Stiftung, Düsseldorf (2016).
19. Sanchez, R., Mahoney, J.T.: Modularity and economic organization: concepts, theory, observations, and predictions. Univ. of Illinois, Illinois (2012).
20. Dubois, A., Gadde, L.-E.: Systematic combining: an abductive approach to case research. J. Bus. Res. 55, 553–560 (2002).
21. Marshall, M.N.: Sampling for qualitative research. Fam. Pract. 13, 522–526 (1996).
22. Gibbert, M., Ruigrok, W., Wicki, B.: What passes as a rigorous case study? Strateg. Manag. J. 29, 1465–1474 (2008).
23. Tuunanen, T., Bask, A., Merisalo-Rantanen, H.: Typology for Modular Service Design: Review of Literature. Int. J. Serv. Sci. Manag. Eng. Technol. 3, 99–112 (2012).

Towards a Cost-Benefit-Analysis of Data-Driven Business Models

Andreas Zolnowski[1], Jürgen Anke[2], and Jan Gudat[3]

[1] Universität Hamburg, Germany
andreas.zolnowski@uni-hamburg.de
[2] Hochschule für Telekommunikation Leipzig, Germany
anke@hft-leipzig.de
[3] Capgemini Deutschland Holding GmbH, Germany
jan.gudat@capgemini.com

Abstract. The emergence of data-driven business models calls for their systematic design and evaluation. In this paper, we focus on a first step towards a Cost-Benefit-Analysis of data-driven business models. Within data-driven business models, data act as enabler for the development of innovative services. However, to justify internal funding of new services, an assessment of the financial impact for the service at hand is often required. We approach this by identifying drivers of cost and benefit based on the Service Business Model Canvases of twenty cases. Based on the results, all drivers and their associated models for quantification were consolidated into a single meta-model. With this, we provide a basis for the economic assessment of service ideas and their refinement during the design process.

Keywords: Data-driven Business Models, Profitability, Service Engineering, Cost-Benefit-Analysis, Smart Services

1 Introduction

Many services have been proposed to transform product-oriented into service-oriented businesses [1, 2]. To systematically design and communicate ideas for new service business models, the Service Business Model Canvas (SBMC) has been proposed by Zolnowski [3]. With the ongoing digitization of service delivery processes, a new class of data-driven services has emerged [4]. While the SBMC is not restricted to data-driven services, it obviously can be used for this purpose as well.

One of the key characteristics of component based business model representations [5], like the Business Model Canvas (BMC) [6] or the SBMC [3], is their qualitative nature, which is very suitable for developing and refining the business logic of an idea [3, 7]. However, to justify internal funding of service engineering initiatives, an assessment of the financial impact for the planned service is required. To this end, various cost items for service provision have to be considered as well as savings through process improvements and revenues through additional offers to customers. Assessing these factors in early stages is challenging but helpful for decision making.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Zolnowski, A.; Anke, J.; Gudat, J. (2017): Towards a Cost-Benefit-Analysis of Data-Driven Business Models, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 181-195

In this paper, we develop a concept for the quantification of data-driven business models (DDBMs). For this, we provide a basis for the economic assessment of service ideas and their refinement during the design process. Hence, we answer the question "What factors need to be considered in a Cost-Benefit-Analysis of a data-driven business model?" To answer this question, we analyze twenty case studies to identify factors that are relevant to describe and evaluate business models. Based on the results, all factors and their associated models for quantification were consolidated into a single meta-model. With this, we strive to improve the overall service design process in practice.

Within this contribution, we focus on a refinement and assessment of business model ideas but do not address the issues of finding a good value proposition, price model or partners. However, with our integrated model, we can support showing the (financial) impact of the design decisions made in the SBMC.

Furthermore, we acknowledge the relevance and importance of non-financial benefits [8-10], especially for innovative offers such as data-driven services. However, in this paper we focus explicitly on the financial dependencies in a DDBM.

The paper is structured as follows. Firstly, we introduce our conceptual foundations with regard to a DDBM. Then we explain the research methodology and introduce the applied case studies. This is followed by the results of the analysis of the cases. Thereupon, a cost-benefit analysis model for a DDBM is proposed and discussed. The paper ends with a conclusion and an outlook.

2 Conceptual Foundation

2.1 Service Business Models

Considering business models, there is a variety of different understandings and definitions [3, 7, 11]. Due to the lack of definitional clarity, alternative conceptualizations of business models exist (e.g. [12-14]) that result in conceptual diversity like a variety of ontologies and representations. Common ontologies for business models are e3-value Ontology [15] and the Business Model Ontology [16]. Representations can be distinguished in two research streams. The first research stream comprises a more flow-oriented perspective on business models. A prominent example for this stream is the e3-Value method [15]. The second research stream comprises a system-level holistic view on the business logic of an economic entity or offering [7]. The most prominent example for this stream is the (BMC) [6].

Fostered by a service based change in value creation [17, 18], business models are also discussed in service research [3, 19, 20]. Service business models are different from product based business models because of the specific characteristics of service. In general, service is a process between interacting parties for the benefit of another party. Especially, the interaction is of high relevance. Known as value co-creation, it is one key aspect of service [18, 21]. Additionally, service value has a unique and phenomenological character [18, 22]. Furthermore, the interaction of service results in a mutual integration of resources and activities. Possible resources that have to be integrated are e.g. skills, knowledge, physical resources and decisions [23, 24].

Because of their specific characteristics, representations for service business models differ from representations for traditional business models [3, 25]. One service specific business model representation is the SBMC. The SBMC highlights the integration of different actors within a service business model and thus, allows focusing on the strategic relevant co-creation in the business logic of service-based business models. As overall logic, the SMBC focuses on the contribution to and benefit of each actor. This logic is applied in the seven dimensions value proposition, relationship, channels, revenue stream, key resources, key activities, and cost structure. In the dimensions customer and key partners, the different actors are defined [3]. The SBMC is displayed in **Figure 1**.

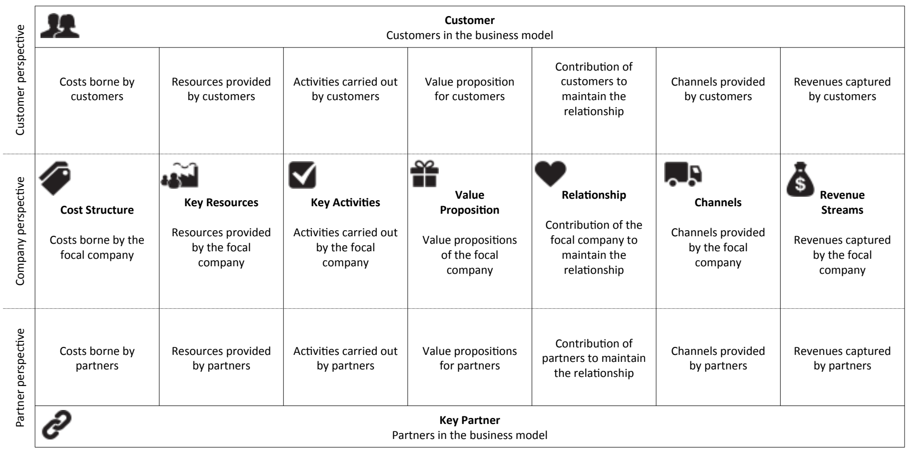


Figure 1: Service Business Model Canvas [3]

2.2 Data-driven Services

Based on the service oriented paradigm new services like data-as-a-service or analytics-as-a-service emerge [26, 27]. Within DDBM, data act as enabler of such innovative services. With enabling technologies, like sensor technology and cloud computing, companies can exploit data from and about their customers. In their own environment companies get enabled to generate new profitable know-how based services [28, 29].

Requiring such new technologies Veit et al. [30] state that "a business model is digital if changes in digital technologies trigger fundamental changes in the way business is carried out and revenues are generated." The BITKOM [31] quotes in their report a similar definition whereupon business models are digital if changes of digital technologies do have fundamental consequences for the business processes and the revenues of the company.

According to Hartmann et al., DDBM is defined as "a business model that relies on data as a key resource" [26]. Brownlow et al. [32] similarly state that "data is obviously fundamental to a DDBM" and Bulger et al. [33] agree, that "data should be central to the business." These definitions are rather simple and differentiate business models on their use of data or not. A more complex perspective on DDBM is proposed by Schüritz and Satzger [4]. According to this, there is no DDBM per se; rather, there is a continuum of options how to provide data-driven service. Hence, there is a smooth transition between business models that use little data and those that enrich all areas of its business model with data analysis [4]. Thereby existing data or new data can be used to either create new business models or enhance existing ones [31]. In the latter case either the value creation, the value proposition or the value capturing or combinations of these can be enhanced by data [4]. For the purpose of our research, a complex differentiation of DDBM is not necessary. Thus, we chose to define DDBM according to Hartmann et al. [26].

The most relevant aspect during the design of a DDBM is the value that should be attained by the data analysis. Hence, the why and how [33] need to be examined. This includes defining the used data. For this purpose, Mathis und Köbler [34] developed a data canvas. The canvas distinguishes between batch and stream as well as internal and external data. Internal stream data do provide the most value since they allow a constant monetization and the data are accessible at any time without any restrictions" [34]. However, to be able to exploit the analysis potential data needs to be well integrated into the business model [33].

2.3 Cost-Benefit Analysis

To make the decision if to invest in a project, a Cost-Benefit-Analysis (CBA) can be performed [35, 36]. A CBA is an established tool for assessing the economic benefit of an investment. As such, it can support decision making on whether a service provider should proceed with the engineering and implementation of a new data-driven business model or not [35, 36]. Due to the complexity of service [37], a systematic capture and analysis of CBA-related factors is a desirable goal. To facilitate end-to-end engineering of smart services, we propose to enrich the qualitative perspective of component based business model representations, in particular the SBMC, with quantitative information. Existing work dealing with business models do not provide means for quantification. Moreover, currently profitability modelling does not include the customer perspective in detail.

One example of a method for assessing data-driven services for connected products was proposed by Anke and Krenge [38]. They propose a meta-model for "Smart Services" from which a business case is derived during the modelling process. In their work "Smart Services" are understood as digitally provided services for connected products. Therefore, "Smart Services" are data-driven services that rely on data which is at least partly provided from connected products, i.e. the Internet of Things. While the meta-model of Anke and Krenge is not directly related to a business model, it provides a connection between service design and its financial evaluation. We therefore will use it as foundation for the concept presented in this paper, as we consider the data-driven business model that might consume data from sensors and the Internet of Things as well.

3 Methodology

In order to design a framework for early-stage profitability assessment, this research applies the following method: first, we identify quantifiable influence factors from a consolidated list of influence factors of digital services on service business models. This list was gained through a multiple case study [39], conducted by Zolnowski et al. [40]. The focus was placed on international companies that successfully developed and implemented successfully data-driven innovations.

The identified cases cover (1) the improvement of the customer orientation, (2) process optimization, (3) optimization of resource consumption, and (4) the collection of information to complement and accelerate decisions. In sum, twenty cases from seven industries were selected and analyzed. Thirteen cases were identified from data of a consulting company and seven cases were derived from literature and public information. The chosen cases cover data-driven innovation projects in different industries (see **Table 1**).

Table 1. Description of the analyzed cases

Companies	*Company description & Examples for implemented data-driven innovation projects*
Three automotive companies	**Two** German automotive manufacturers >70,000 employees (2014) **One** from the automotive parts industry >30,000 employees (2014)
	Project: Predictive Maintenance by expansion of sensors on assets; Optimization of processes by data integration
Seven manufacturing companies	**Three** German companies >6,000 employees **Two** German companies >63,000 employees **Two** American companies >80,000 employees
	Project: Predictive Maintenance by expansion of sensors on assets; Service innovation and use of Internet of Things; Optimization of processes by data-driven forecasting
Five logistics and transportation companies	**One** joint venture, 51-200 employees (2014) **Four** companies, 1800-5.000 employees (2014)
	Project: Coordination of infrastructure by real time data of players; Tracking of assets by expansion of sensors
Two retail companies	**One** German retail company >17,000 employees (2014) **One** Swiss food company >300,000 employees (2013)
	Project: Optimization of disposition by analysis of market data
One insurance Company	**One** American start-up, 201-500 employees
	Project: Product innovation in car insurance by use of Internet of Things
One energy company	**One** German electric utility company > 50,000 employees (2014)
	Project: Predictive Maintenance by expansion of sensors on assets
One telecommunication Company	**One** Swiss telecommunication provider >20,000 employees
	Project: Coordination of infrastructure by data of passenger traffic

The results of this work are shown in **Figure 2**. This figure shows the identified influence of data-driven innovation projects on the business models of the analyzed companies. The effects are symbolically illustrated by gray boxes and grouped if being similar with bold titles describing aggregated types of effects. Because of the networked character of a DDBM, all influences are differentiated according to their impact on customer, company, or partner. Thus, elements that have a direct influence on the customer, are classified to the customer perspective.

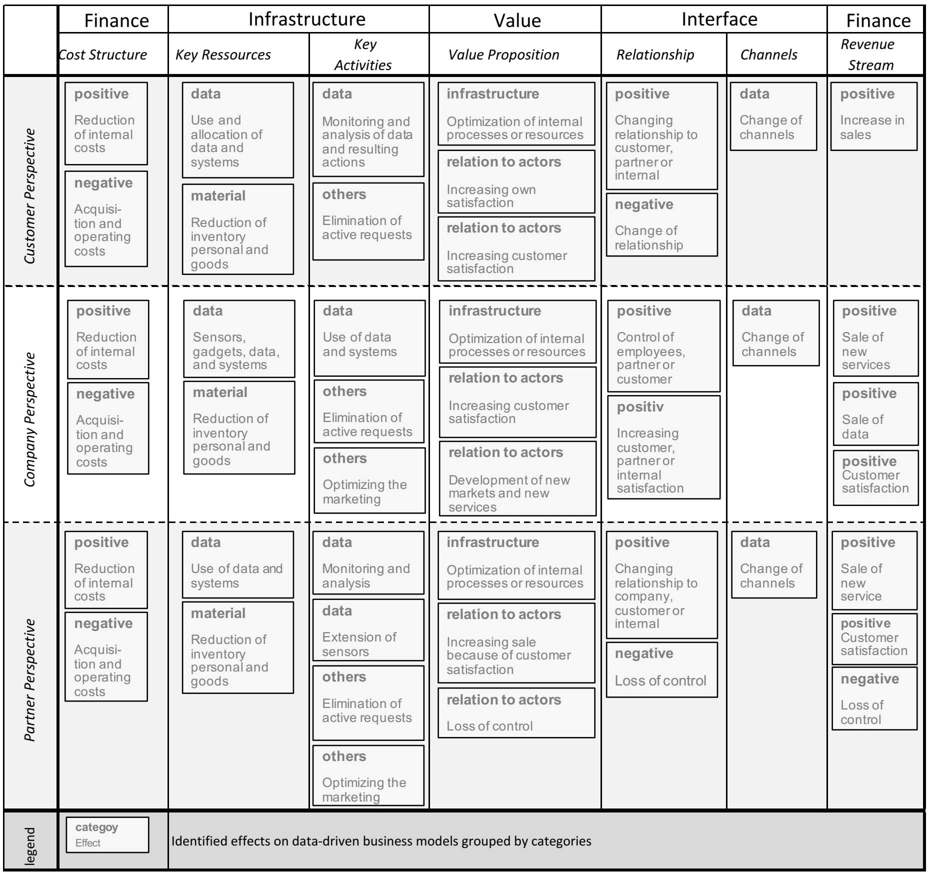


Figure 2: Identified effects of data-driven innovation projects [40]

In the next step, we analyzed the identified influence factors according to their quantitative or qualitative nature. This was necessary to identify those factors that have a quantitative influence on the costs and benefit of the DDBM. Based on this information, we were able to determine the influence factors for a CBA of a DDBM on an empirical basis. To summarize, the proposed CBA model adapted for the DDBM is developed in the following steps:

1. Categorization of quantitative influence factors from twenty case studies into cost, revenue and savings.
2. Development of a parameter set to simplify the capture of relevant inputs and derive the financial values for identified influence factors.
3. Integration of the parameter sets into an integrated meta-model, based on the meta-model proposed by Anke and Krenge [38].

The remainder of this paper is structured as follows: In section 4, we conduct step 1 and 2, while section 5 covers step 3 followed by a discussion of the results. The paper concludes with an outlook on further research questions.

4 Case Analysis

4.1 Identification and classification of CBA-related parameters

In our multiple case analysis, we were able to differentiate between qualitative and quantitative influence factors on the business models. Qualitative factors comprise effects like increasing customer satisfaction or change of relationship that cannot be translated directly to a countable metric. However, influence factors with a quantitative nature enable a direct analysis of countable and monetary consequences. These factors comprise effects like reduction of internal costs or sensors, gadgets, data, and systems. As already stated, we acknowledge the relevance and importance of non-financial benefits [8-10]. Nevertheless, in this paper we focus on quantitative influence factors and exclude qualitative influence factors intentionally. Within these factors we are able to distinguish between three classes of effects. In the following, we present the results of our multiple case analysis according to the identified classes (1) costs, (2) revenues, and (3) savings.

4.2 Costs

There are a variety of costs that can be directly related to the development and management of a DDBM. This includes all necessary preparation and the use of data and systems (including the influence factors different use of data and systems; monitoring and analysis of data, actions; and extension of sensors in the key activities and use and allocation of data and systems; sensors, gadgets, data, and systems; use of data and systems in the key resources). Especially in the development phase, the improvement or implementation of infrastructure is an important cost factor. For example, in a manufacturing case, customers have to implement remote services hardware in their machines to collect data and enable the connectivity between the customer's machines and the provider's servers. A driving force for the improvement or implementation of new infrastructure is the lack of sensors, actuators, and connectivity in older machines. Highly depending on the industry, these elements can already be part of an existing infrastructure or they need to be added. However, all three elements are enabling technologies for the DDBM. Sensors are necessary to monitor machines and collect data. Connectivity establishes a link between the systems of the customer and provider to transmit the collected data, and actuators enable the provider to remotely take control or even change things automatically based on data. If relevant infrastructure is missing, companies having no or limited experience and are often surprised about the advancement in sensor technology. Off-the-shelf products can cost a few cents per sensor. Even equipping an existing product with an additional sensor can result in marginal extra costs. However, if customized sensors are required, development costs can reach several hundred thousand Euros.

Besides the infrastructure, specialized software can be necessary for a DDBM. As our cases show, such software can be developed, purchased, or leased. Alternatively, cloud services can be applied. According to the respective decision, in many cases the costs occur on-demand, recurring, or for the development of the software. Integrated in this software, algorithms enable the processing of the data. These algorithms can be highly individualized and need to be developed, maintained, and processed by the provider or other partners. This also applies to the resulting reports of this process.

To enable the whole DDBM, the connectivity between all actors and their infrastructure and software is of high relevance. According to the type of connectivity, e.g. permanently or recurring, the actors have to calculate with different pricing models.

According to our analysis, the influence of infrastructure, software, and connectivity have an important influence, with a direct effect on the focal company, the customer, and partner. Hence, in order to introduce a DDBM, invests into all factors are needed in the entire service system.

4.3 Revenues

As our cases show, in a DDBM, revenues can be enabled for any actor. From a company perspective, revenues can be generated from sales of new services and possibly also from the sale of data to third parties. Despite an existing relation to a customer, an increase in revenue is not mandatory with existing customer satisfaction. In our cases we were able to identify companies that establish completely new DDBMs that were offered to the customers. For example, a manufacturer facilitates higher safety standards by offering the tracking of tools in the maintenance process of the customer. This leads to an increase of responsibility through employees and avoids occurrence of abandoned tools.

But also customers and partners are enabled to generate additional revenues. By facilitating existing or enabling new processes, they can improve their existing or establish completely new business models. Target of these operations is to increase sales or to exploit economies of scale.

4.4 Savings

Beside of revenues, another positive economic effect are savings. In particular, the implementation of a DDBM allows for optimization of processes or reduction of assets, which both lead to lower costs by the reduction of inventory (key resources).

From a company's, partner's, and customer's perspective, DDBMs have direct influence on the operational processes. The processes can be optimized in a different manner. In one case, the provider gathered data about his and the customer's processes. Based on his knowledge, he was able to provide consulting services to his customer and to optimize his processes. Another case illustrated the elimination of active requests. Within this case, manual requests and process executions were replaced by automatized processes. This led to a reduction of operating costs.

Additionally, a reduction of inventory, personnel, and goods is possible. In particular, we observed a reduction of resources by an optimization of resource planning and hence, adjusted resource utilization.

5 A Cost-Benefit-Analysis Model for DDBM

5.1 Parameter Categories and financial quantification

In general, a CBA considers all costs and benefits to assess the economic value of an investment project. Cost refers to the financial effort required to build and operate a DDBM system. For all effects created by the investment project at hand, monetary values have to be assigned. Benefits can be either additional revenue or cost savings, e.g. through improved process efficiency. In the context of business models, we expect a CBA to support in experimentation and finding better solutions through an additional evaluation criterion. As a foundation for the CBA model, we use a series of payments, which contains the net cash effect per planning period.

As we take the perspective of the service provider, we build our model based on the quantitative factors of the focal firm identified in the previous section. The required elements and their relationships will be expressed as a UML class-diagram meta-model, which is why we use the terms class and attributes in the description below. The complete meta-model is shown in the next section.

Cost is already expressed in monetary values. However, it is usually difficult to estimate a total value. Therefore, we propose to break down cost into various more concrete items, which can be estimated with higher confidence. As we have identified cost items that are relevant for a DDBM, this helps to create a CBA model for these scenarios. We differentiate between one-time, recurring constant, and recurring-growing cost. For simplification, growth rates are fixed per period (see Table 2).

Table 2: Translation of costs into the capture model

SBMC Factor	*Capture Model*
Acquisition and operating costs	Operating costs are recurring, and can be both constant and growing. We use the `OperationCost` class to describe cost of various `CostTypes` and `PaymentIntervals`. For all costs related to the use of external services, such as weather info, messaging etc., we provide the `ServiceVariableCost` class. It relates to the `Functions` and their usage of `ExternalService`. `Functions` can also use `DataPoints` from connected `Devices`. As the latter are provided from connected products, these can be modeled as `Device` with attributes for `costPerMBTransfer`, `devicePrice`, `initialDeployment` and `growthPerYear`.
Sensors, gadgets, data and systems	Equipment of all kind has to be purchased (one-time cost), so type of equipment, price and quantity capture these cost. In the model, these are captured in the `InitialInvest` class.

Revenue is created by providing value to customers. It is also already expressed in monetary values. As with costs, we propose to break down revenue into more concrete offers with a single price. This helps decision makers to describe the service in a more specific way and see the impact of various configurations (see Table 3).

Table 3: Translation of revenues into the capture model

SBMC Factor	*Capture Model*
Sale of new services	An `Offer` can be modelled with a `offerName` and `price`, which can be interpreted as subscription or transaction-based price. The quantity is defined using the `CustomerDemand` class, which contains attributes for `customerGroupName`, `initialYear` and optional `growthRatePerYear`.
Sale of data	

Savings refer to reduction of cost at the service provider. They can be created through process efficiencies, reduction of stock, resource consumption etc. To quantify these effects, the internal organization of the service provider has to be known in great detail. However, this would increase the complexity of the CBA model greatly. Therefore, we propose to model savings as relative improvement to a certain level, which can be expressed by three simple parameters. In the following table, we list the identified SBMC parameters by category, and show how we quantify them. Please note, that the factor "Reduction of internal costs" is not explicitly mentioned, as it is covered by the four factors listed in the table (see Table 4).

Table 4: Translation of savings into the capture model

SBMC Factor	*Capture Model*
Reduction of inventory, personnel and goods	To quantify these effects, we use the Savings class, which can be instantiated for every relevant savings effect. It is modelled with the name of the factor, its `initialLevel` as monetary value and the `reductionInPercent` to capture the savings effect.
Elimination of active requests	
Optimizing the marketing	
Optimization of internal processes or resources	

5.2 Integrated Meta-Model

All factors and their associated models for quantification were consolidated into a single meta-model, which is depicted as a UML class diagram. Class diagrams are an established way of representing the structure of domains semi-formally. Classes represent entities and their attributes. Relations are expressed with associations, which can also be qualified with cardinalities to show how many instances of one class can be related to a certain number of the associated class.

Our integrated meta-model is depicted in **Figure 3** below. It adapts and extends the smart service and business case model proposed by Anke and Krenge [38]. All white classes are part of the quantification model described in the section above. The classes with grey a background are used to describe the basis for calculating the CBA. From the original model, we mainly reused the parts concerning the `DataPoints` from `Devices`, their usage in `Functions` which are subsequently bundled in `Offers`. Furthermore, the concept of `ExternalServices`, their usage as well as `Projects` were part of the original model. Our extensions are mainly related to `Savings`, the `SMBC_Factor` relation and `InitialInvest`. Also, the modelling of cost was extended to enhance flexibility.

The starting point is a `Project`, which contains a number of `planningYears`, a `name` and a derived attribute `financialResult`. It is calculated from a series of payments, which is represented by `ProjectYears`, which in turn contain derived attributes for `revenue`, `cost` and `savings`. Costs can either be manually specified `OperationCost` or automatically calculated using the class `ServiceVariableCost` for data transfer and external services. This only applies to services (Offers), which are modeled using `Functions`, `DataPoints`, `Devices` and `ExternalServices`.

For a concrete project, the meta-model has to be instantiated. The service designers start with a qualitative design of the DDBM in a SBMC. Each factor in the SMBC can then be represented by either an Offer, Savings or Cost item. The quantification of each item is achieved through the parameters described in section 5.1.

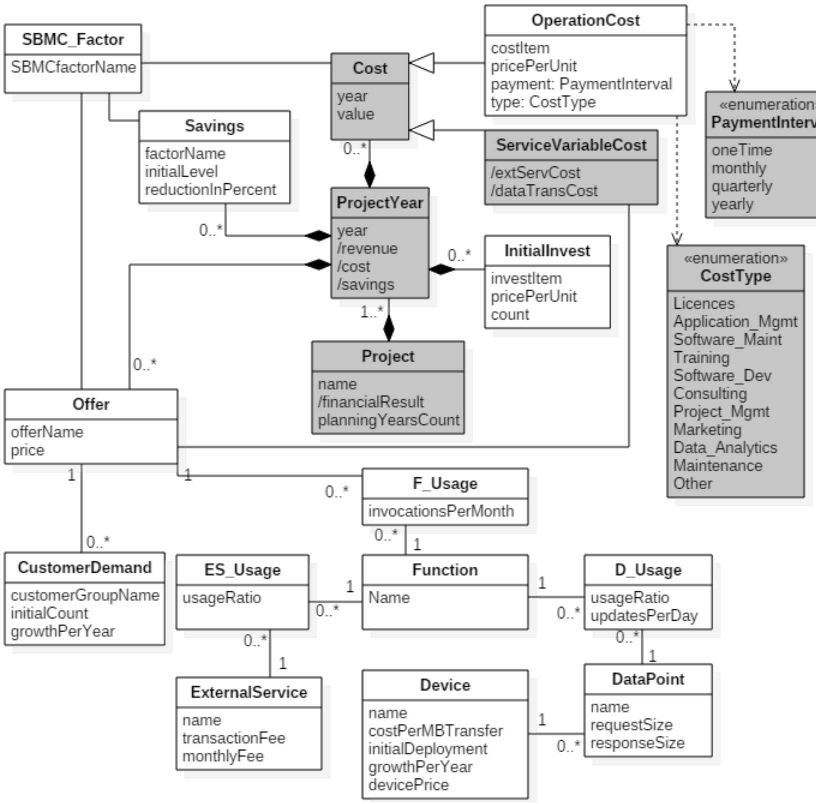


Figure 3: Integrated Meta-Model

5.3 Discussion

The high complexity and qualitative nature of service make it difficult to judge the financial impact of services, like of a DDBM, in early conception stages. While methods for financial decision making, like Net Present Value or Return on Investment, are well established, they are rather generic and not linked to the specific development of business models. Our proposed meta-model establishes links between the qualitative dimensions of the SBMC and main drivers of business value in a DDBM, i.e. new offers, savings and associated cost. To support the refinement of a SBMC, the model includes parameters (e.g. prices, costs and quantities), which are sufficiently detailed to allow for a first estimation during the design process. As these factors were derived from case studies focusing on data-driven innovation projects, our meta-model is designed to facilitate the development of a DDBM. This is expressed through dedicated elements that must be considered during the development.

In summary, an integrated meta-model divided into three states was derived. To apply this meta-model, firstly, a user has to develop a business idea and fill out a SBMC. Based on this information, a refinement of the business model with Cost-Benefit-Analysis related parameters has to be conducted. To allow an assessment of a DDBM innovation project, we propose concrete influence factors (see `CostTypes` and `PaymentIntervals`) that have to be considered in its specific CBA. In addition to the costs, we were able to derive potential savings and revenues in DDBM initiatives. Finally, a business case can be derived and decision can be taken, whether to proceed with the implementation of this service or not. To improve the creation and refinement of models as well to perform calculations, a software tool can be of great benefit. We see the development of the meta-model as a starting point to develop such a tool in the future.

6 Conclusion and Outlook

In this paper we propose a first step towards a Cost-Benefit-Analysis of data-driven business models and therewith address an important issue of companies in the field of existing service development and service engineering initiatives [35, 36]. We analyzed twenty case studies on data-driven innovation projects and derived influence factors that have an impact on a business model. A set of parameters was developed that allow identifying relevant inputs and deriving financial values for the included factors. As an extension of the existing smart service model [38], an integrated meta-model was derived that if being applied allows for the quantitative evaluation of a DDBM. This method enables decision makers to evaluate and calculate their business case for a data-driven innovation through refinement of a business model.

Our integrated meta-model provides a theoretical contribution as it helps researchers by fostering the understanding of financial dependencies in data-driven innovation processes towards new business models. By analyzing data-driven innovation projects, we were able to determine quantitative influence factors that have a direct monetary impact on a DDBM. Based on this knowledge, it is possible to create, shape, and improve tools and methods that foster service innovation and the design of a new DDBM. Practitioners can utilize these results in order to foster the development of data-driven innovations in their servitization efforts. Moreover, they can analyze different innovation projects in regard to their financial effects and thus, better intercept business potentials.

Nevertheless, also some limitations have to be considered. Firstly, this paper focuses on quantitative influence factors in the development of a new DDBM. This decision was made purposely and need to be addressed in further research. Hence, qualitative influence factors must be considered and their impact added to the meta-model. This includes qualitative criteria, such as improved customer satisfaction, loss of control, competitive advantage. The twenty case studies on data-driven innovation projects analyzed in this paper could be extended with new cases in further research.

Besides addressing these limitations, further research should focus on the practical application in a field experiment and/or lab experiment to evaluate the meta-model and its benefits. Equations need to be developed allowing the computation of the overall benefit. Subsequently, instructions describing how to apply the meta-model in a concrete scenario could be developed.

References

1. Neely, A.: The Servitization of Manufacturing: An Analysis of Global Trends. In: 14th European Operations Management Association Conference (2007)
2. Baines, T.S., Lightfoot, H.W., Benedettini, O., Kay, J.M.: The servitization of manufacturing: A review of literature and reflection on future challenges. Journal of Manufacturing Technology Management 20, 547-567 (2009)
3. Zolnowski, A.: Analysis and Design of Service Business Models. University of Hamburg, Hamburg (2015)
4. Schüritz, R., Satzger, G.: Patterns of Data-Infused Business Model Innovation. IEEE Conference on Business Informatics (CBI), Paris, France (2016)
5. Beha, F., Göritz, A., Schildhauer, T.: Business Model Innovation: the Role of Different Types of Visualizations. In: ISPIM Conference Proceedings. The International Society for Professional Innovation Management (ISPIM), (2015)
6. Osterwalder, A., Pigneur, Y.: Business Model Generation. John Wiley & Sons, Hoboken (2010)
7. Zott, C., Amit, R., Massa, L.: The Business Model: Theoretical Roots, Recent Development, and Future Research. Journal of Management 37, 1019-1042 (2011)
8. Vargo, S.L., Lusch, R.F.: Institutions and axioms: an extension and update of service-dominant logic. Journal of the Academy of Marketing Science 1-19 (2015)
9. Chandler, J.D., Vargo, S.L.: Contextualization and value-in-context: How context frames exchange. Marketing Theory 11, 35-35 (2011)
10. Edvardsson, B., Gustafsson, A., Roos, I.: Service portraits in service research: a critical review. International Journal of Service Industry Management 16, 107-121 (2005)
11. Fielt, E.: Business service management: understanding business models. Whitepaper, Smart Services CRC (2011)
12. Afuah, A., Tucci, C.L.: Internet Business Models and Strategies. Text and Cases. Mcgraw-Hill Higher Education (2001)
13. Al-Debei, M.M.: The design and engineering of innovative mobile data services: An ontological Framework founded on business model thinking. Brunel University, London (2010)
14. Zott, C., Amit, R.: Business model design and the performance of entrepreneurial firms. Organization Science 18, 181-199 (2007)
15. Gordijn, J.: Value-based Requirements Engineering-Exploring Innovative e-Commerce Ideas. Vrije Universiteit, Amsterdam, NL (2002)
16. Osterwalder, A.: The Business Model Ontology - a proposition in a design science approach (2004)
17. Grönroos, C.: Adopting a service business logic in relational business-to-business marketing: value creation, interaction and joint value co-creation. pp. 269-287 (2008)
18. Vargo, S.L., Lusch, R.F.: Service-Dominant Logic - Premises, Perspectives, Possibilities. Cambridge University Press, Cambridge (2014)
19. Bouwman, H., Fielt, E.: Service Innovation and Business Models. In: Bouwman, H., De Vos, H., Haaker, T. (eds.) Mobile Service Innovation and Business Models, pp. 9-30. Springer Berlin Heidelberg (2008)

20. Fielt, E.: A ‘service logic’ rationale for business model innovation. EURAM Annual Conference 2012, Rotterdam (2012)
21. Grönroos, C.: Conceptualising value co-creation: A journey to the 1970s and back to the future. Journal of Marketing Management 28, 1-15 (2012)
22. Edvardsson, B., Tronvoll, B., Gruber, T.: Expanding understanding of service exchange and value co-creation: a social construction approach. Journal of the Academy of Marketing Science 39, 327-339 (2010)
23. Grönroos, C., Ravald, A.: Service as business logic: implications for value creation and marketing. Journal of Service Management 22, 5-22 (2011)
24. Moeller, S.: Customer Integration - A Key to an Implementation Perspective of Service Provision. Journal of Service Research 11, 197-210 (2008)
25. Ojasalo, K., Ojasalo, J.: Adapting business model thinking to service logic: an empirical study on developing a service design tool. THE NORDIC SCHOOL 309 (2015)
26. Hartmann, P.M., Zaki, M., Feldmann, N., Neely, A.: Big data for big business? A taxonomy of data-driven business models used by start-up firms. A Taxonomy of Data-Driven Business Models Used by Start-Up Firms (2014)
27. Chen, Y., Kreulen, J., Campbell, M., Abrams, C.: Analytics Eco-system Transformation: A Force for Business Model Innovation. Annual SRII Global Conference (SRII), pp. 11-20, San Jose, CA, USA (2011)
28. Zolnowski, A., Böhmann, T.: Veränderungstreiber service-orientierter Geschäftsmodelle. In: Böhmann, T., Warg, M., Weiß, P. (eds.) Service-orientierte Geschäftsmodelle. Springer-Verlag, Berlin Heidelberg (2013)
29. Engel, T., Sadovskyi, O., Boehm, M., Heininger, R., Krcmar, H.: A Conceptual Approach for Optimizing Distribution Logistics using Big Data. Twentieth Americas Conference on Information Systems (AMCIS 2014), Savannah (2014)
30. Veit, D., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Kundisch, D., Leimeister, J., Loos, P., Spann, M.: Business Models - An Information Systems Research Agenda. WIRTSCHAFTSINFORMATIK 56, 55-64 (2014)
31. BITKOM: Big Data und Geschäftsmodell-Innovationen in der Praxis: 40+ Beispiele. (2015)
32. Brownlow, J., Zaki, M., Neely, A., Urmetzer, F.: Data and Analytics-Data-Driven Business Models: A Blueprint for Innovation. Cambridge Service Alliance (2015)
33. Bulger, M., Taylor, G., Schroeder, R.: Data-Driven Business Models: Challenges and Opportunities of Big Data. Oxford Internet Institute (2014)
34. Mathis, K., Köbler, F.: Data Canvas und Data-Need Fit. Mensch und Computer 2015–Usability Professionals (2015)
35. Boardman, A., Greenberg, D., Vining, A., Weimer, D.: Cost-Benefit Analysis. Pearson Education, New Jersey (2011)
36. Edward J., M., Quah, E.: Cost-benefit analysis. Routledge, New York (2007)
37. Böhmann, T., Leimeister, J.M., Möslein, K.: Service Systems Engineering. Business & Information Systems Engineering 6, 73-79 (2014)
38. Anke, J., Krenge, J.: Prototyp eines Tools zur Abschätzung der Wirtschaftlichkeit von Smart Services für vernetzte Produkte. MKWI 2016, Ilmenau, Germany (2016)
39. Yin, R.K.: Case study research: Design and methods. Sage publications (2014)
40. Zolnowski, A., Christiansen, T., Gudat, J.: Business Model Transformation Patterns of Data-Driven Innovations. European Conference on Information Systems (ECIS 2016), Istanbul (2016)

Measuring National Culture by Analyzing Business Processes: A Case Study in Germany and India

Benjamin Wehner[1], Thomas Falk[1], Susanne Leist[1], and Christian Ritter[1]

[1] Department of Management Information Systems, University of Regensburg, Germany
{benjamin.wehner, thomas.falk, susanne.leist, christian.ritter}@wiwi.uni-regensburg.de

Abstract. Nowadays, many companies face problems because of cultural differences, especially in multinational settings. Traditionally, national cultures have so far been identified by questionnaires asking participants about e.g., their values. These invisible elements of culture become manifest in tangible artifacts such as concrete actions or structures, e.g., rituals and organizational charts. Process models serve as a graphical representation of processes precisely describing activities, responsibilities and process flows. Thus, we anticipate that the behavior becoming apparent in process models provides insights into national characteristics. Consequently, the goal of this paper is to develop an approach to measure national culture in process models. Based on Hofstede's (2010) cultural dimensions, we define metrics that can be applied to process models. We demonstrate the use of these metrics by applying them to a process executed both in a German and an Indian company. Our analysis confirms a correspondence of the metrics' results with Hofstede's findings.

Keywords: National Culture, Business Process, Measurement Approach, Case Study

1 Introduction

National cultural characteristics and their influence on collaboration, both within and among companies, is an important research area in times of globalization [1,2]. Neglecting cultural peculiarities may lead to disastrous consequences. A plane crash involving Avianca Airlines in the 1990's, which was, ultimately, caused by a specific cultural behavior of the crew members may serve as a warning example. In this crash, the co-pilot did not challenge a wrong interpretation of the instruments by the flight captain because he came from a country where subordinates are generally afraid of expressing disagreement with superiors [3]. In IS research, problems caused by cultural differences have been addressed, too. For example caused by different communication styles (in Japan and the US) leading to poor information sharing or caused by different norms for decision-making (Mexico and South Korea) [4].

To make a global collaboration setting more beneficial for all parties, it is essential to be aware of all cultural peculiarities and differences of the individual stakeholders, e.g., customers and suppliers in foreign countries [2]. As all of these stakeholders are

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Wehner, B.; Falk, T.; Leist, S.; Ritter, C. (2017): Measuring National Culture by Analyzing Business Processes: A Case Study in Germany and India, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 196-210

connected by business processes, the analysis of cultural peculiarities in relation to business processes is a promising research area. Schein [5] highlights that "*for purposes of cultural analysis, the organizational processes by which such behavior is made routine*" (p. 26) can be used. Business processes are generally captured by process models in various process modeling notations, all of them aiming at graphically representing the actual process as closely as possible. Thus, we anticipate that different cultural aspects can be identified in these process models, as they reveal the behavior, responsibilities, and the sequence flow of tasks.

To date, research has assessed national culture by the use of questionnaires (e.g., [6,7]), asking respondents about their beliefs and assumptions. As questionnaires are very time-intensive, expensive and the intended behavior is only analyzed, a measurement approach that reflects observable behavior on the basis of the employees' daily routines complements questionnaires with new aspects. For example, details on the hierarchy in a company can be inferred [8] when analyzing who makes decisions in a process model. In this paper, we analyze process models by means of metrics, as metrics are a well-known means to identify influencing factors on processes [9]. We anticipate that an analysis of business processes in terms of culture will provide a company with valuable insights into the cultural attitude of the process participants, e.g., if employees culturally fit to the tasks they perform. This knowledge can then contribute to various phases of the BPM lifecycle: To design a process that is perfectly aligned to the cultural attitude of the process participants, to determine a cultural fit of two companies' processes during mergers and acquisitions [10] or to analyze if culture is the origin of process performance problems, a correlation that has recently been identified [11]. Altogether, the goal of this paper is to develop a measurement approach applied to business process models, which in a first step measures selected national characteristics. The resulting metrics form the basis for discussions about discrepancies between national culture and its manifestation in companies.

The remainder of this paper is organized as follows: in section 2, we provide an overview of culture, business process management and measurement of national culture. Section 3 presents the operationalization of the cultural dimensions, the deduction of metrics, and the process of data collection. In section 4, we apply these metrics to process models of two companies, one from Germany and one from India, and present the results. The insights gained from the application of the metrics are discussed by comparing them to the comprehensive study by Hofstede [6] in section 5. The last section gives a summary of our research, names limitations and proposes options for further research.

2 Conceptional Background

2.1 Culture

Culture is a rather diffuse concept with different meanings depending on the context [12]. A definition often used, as it comprises the general aspects of culture, was evolved by Schein [5] with culture being referred to as shared values of a group, which can be recognized in actions and structures.

Culture can generally be presented as an iceberg model [13], with both visible and invisible elements. Invisible elements are *Underlying Assumptions* (e.g., ideology, feelings, taken-for-granted beliefs [13,14]) as well as *Espoused Beliefs and Values* (e.g., morale, ethical norms) [1,5], often subsumed by researchers under the term values [14,8]. Invisible elements of culture manifest themselves in tangible *Artifacts* [5], which represent the visible elements such as concrete actions or structures (e.g., behavior, organizational charts) [13]. According to the range of influence, different levels of culture can be defined, e.g., national, organizational, and subgroup culture [1,14]. Each cultural group shares characteristics distinguishing one from another [6].

Basically, the culture in a company is influenced by the deeply embedded national culture as well as the particular organizational culture [10]. However, the organizational culture is always subject to the basic assumptions of the national culture. This has been demonstrated in various cases by Schneider [15], who analyzed human resource practices in multinational companies. In these cases, national characteristics had a stronger influence on work practices than corporate identity. This fact was also confirmed by other surveys (e.g., [16]). Even though there will always be a coherence between organizational and national culture, Hofstede et al. [17] point out that they are also distinctive due to different cultural dimensions (a list of dimensions is presented in [14]).

In this paper, we focus on national culture that is stable over time [2] and difficult to influence as it represents the basic behavior and values of a particular society [1] and has a predominant effect on people as compared to the organizational culture [10,15]. Even though the concrete personal behavior of individuals belonging to a national group may differ, generally valid tendencies can be observed for certain cultural dimensions [6,8].

2.2 Culture and Business Process Management

Culture has been identified as an important aspect of BPM [18,19], too, and the number of publications linking the two topics has indeed increased in recent years. Research in this area can be classified on the basis of two dimensions: the interrelation between culture and BPM and the referenced cultural group. In this respect, the main emphasis lies on organizational culture and its impact on BPM activities or business process performance (cf. [20,21]). Besides, the concept of a distinct *BPM culture* was identified and its values operationalized [22].

The role of national culture in BPM has been much less highlighted in existing research. Authors dealing with this topic analyze the national influence on BPM [23] and the application of BPM concepts in different national contexts [19,24]. Central to BPM is the notion of a business process consisting of a cohesive sequence of functions that create an output by adding value to the input and thus fulfil an organizational task [25]. Business processes represent a socio-technical system in which humans collaborate and carry out the single tasks to achieve the process output [26].

2.3 Measurement of National Culture

So far, national culture in a broader sense has been assessed by means of surveys, asking respondents questions about their feelings, beliefs, morale, or how they would behave in certain situations [6,7]. Well-known examples are a survey by Hofstede [6] with approx. 116,000 respondents in 76 countries and the GLOBE study with 17,300 interviewees in 63 countries [7]. Questions in cultural surveys generally target the two layers of *Underlying Assumptions* and *Espoused Beliefs and Values*. As a result, visible *Artifacts* are not dealt with. An important category of visible *Artifacts* are process models graphically representing business processes using semi-formal modeling techniques. In a process model, the process is decomposed into a set of interrelated activities that are logically and temporally connected. Using semi-formal modeling techniques helps to increase the specificity of the description and to avoid the ambiguity, which, for instance, natural or narrative texts often imply [27]. Therefore, process models precisely describe the interaction among people, technology and organizations with the aim of improving the effectiveness and efficiency of organizations. Whereas to-be process models represent how people should interact from a management or process owner view, our analysis focuses on as-is process models. As we want to capture the real behavior of people and how they execute their tasks and activities, we analyze the process in its current form. Since as-is models describe this behavior, they reflect the culture of individuals in a country, thus being different from models of other nations. This difference is expected to be detectable and can be measured for each national dimension.

3 Measurement Approach

Our approach is based on predefined metrics that are applied on to business process models. The definition of the required metrics comprises three steps: first, we analyze the operationalization of cultural dimensions in general. Second, three national dimensions by Hofstede et al. [8] are presented and then used to deduct metrics to measure them. Finally, we describe the process of data collection, which is the basis for the calculation of the metrics.

3.1 Operationalization of Cultural Dimensions

The best-known classification in national cultural research are the dimensions by Hofstede et al. [8] (cf. [2,14]), which represent the basic behavior and nature of national cultures: Power Distance, Uncertainty Avoidance, Individualism, Masculinity/Femininity, Indulgence, and Long-Term Orientation [8]. Further studies expand these basic dimensions by detailing or extending them, by e.g., assertiveness, performance, and humane orientation [7], time orientation, and locus of control [14]. As Hofstede describes the most popular conceptualization of national culture [28,29], we take his taxonomy as a basis. In this first step, we develop metrics for the dimensions *Power Distance*, *Uncertainty Avoidance* and *Individualism*, as these were identified as fundamentally distinguishable between nations [6] and are very well defined so that it is possible to derive metrics. In the next section, these dimensions are explained in detail.

We follow the approach by Harvey [30] who operationalized the dimension *Uncertainty Avoidance* to assess differences between design documents and actual design practice. He defined particular characteristics of the dimension *Uncertainty Avoidance* and discussed the differences in each characteristic. We transfer and formalize this approach as follows: Generally, each cultural dimension (CD) comprises several characteristics (C) that can be measured. Metrics (M) are defined to measure the extent of each characteristic. Afterwards, the metrics are aggregated (Agg) for each cultural dimension to obtain a tangible value that can further be interpreted and compared. Figure 1 provides an overview of all components of our measurement system (cf. [30]).

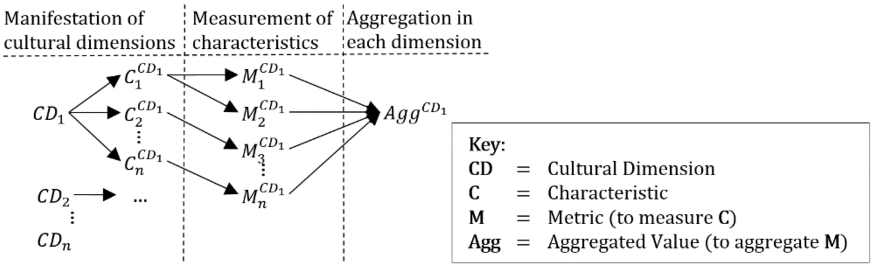


Figure 1. Measurement Approach

To obtain the values for the metrics, the business process models are analyzed. There is a great variety of process modeling notations, the most popular ones being BPMN, EPC and the UML Activity Diagram. Even though different notations exist, most of them offer similar key elements: functions, organizational units, application systems, information objects and connectors [25]. These elements are used to calculate the metrics by analyzing the semantics and the syntax of a process. Elements can be classified with regard to semantics by their element type and by interpreting their labels. Syntax is analyzed to identify sequential, parallel, or alternative process paths (see section 3.3).

3.2 Deduction of Metrics from National Dimensions

The three aforementioned cultural dimensions (CD) by Hofstede et al. [8] are manifested in well-defined characteristics (C). In each dimension, we derived the characteristics analyzing the definitions and descriptions [30]. Furthermore, we analyzed the single items of the questionnaire by Hofstede et al. [8] to stay as close as possible to the original intention of each dimension. As there is also a manual available on how to calculate each dimension on the basis of the questionnaire items, the relevant items in each dimension were selected[1]. Both the definitions and the items in the questionnaire were finally used to derive characteristics that can be developed to metrics, which, in turn, can then be applied to process models. In the following, the deduction of the dimension *Power Distance* is presented in detail and the deduction of the metrics from the other two dimensions is summarized.

[1] The questionnaire and the manual how to use it are available at: http://geerthofstede.com/research-and-vsm/vsm-2013/ (last access: 24.10.2016)

In a cultural group, the power of individuals is not equally distributed. *Power Distance* is defined as the extent of this power and the degree as to which the less powerful accept the existing distribution [6]. In the questionnaire, four questions are asked (item numbers m02, m07, m23, and m26) to calculate the Power Distance Index (PDI) as it is originally named [8]. The questionnaire asks e.g., how important it is for employees to be consulted by their boss in decisions involving their work (m07) and how often subordinates are afraid to contradict their boss (m23). On the basis of the definition and the questionnaire items, two characteristics can be defined: acceptance of unequal distribution of power and unequal distribution of power. With regard to process models, they do not 'betray' any personal feelings, and it is not possible to identify the degree of acceptance of inequality by the less powerful ones in them. Thus, we focus on the extent of unequal distribution of power. The original intention to measure the extent of unequal distribution of power can be identified by looking at single functions, which can be performed by an organizational unit either on a lower or on a higher level. These tasks are checks of results, which is a source of contradiction (m23) and making decisions (m07). If a check is performed by a superior and not by a regular employee, the power is focused on a higher level in that particular case. Thus, the derived metric for a whole process indicates the share of check functions performed by organizational units on a higher level in relation to all check functions performed within the process ($PD_1 = \frac{|CF_{HL}|}{|CF|}$). The assignment of the right to make decisions is the second area for which to show the distribution of power depending on the hierarchy level of organizational units. The metric indicates the share of decision functions that are performed by organizational units on a higher level in relation to all decision functions in the process ($PD_2 = \frac{|DF_{HL}|}{|DF|}$). For both metrics, a high value indicates high *Power Distance*.

Uncertainty Avoidance deals with the unpredictability of situations. It is defined as the extent to which individuals try to avoid uncertain situations by relying on common norms, rituals and practices, e.g., by using standardized documents [6,30]. Additionally, the items m16, m20, m24, and m27 in the questionnaire to calculate the "Uncertainty Avoidance Index" (UAI) are in focus, e.g., the question as to which extent an employee agrees with the statement if one can be a good manager without having precise answers to a question a subordinate may raise (m24). To measure the first characteristic (avoidance of uncertain situations), we explicitly search for functions dealing with quality issues, e.g., the use of a checklist or the four-eyes principle. Quality functions are usually meant to document a current state or to check for a deviation from the outcome, which indeed reduces uncertainty in the specified topic. Thus, if the share of quality functions in relation to all functions is high, a high degree of uncertainty avoidance is measured. In addition, information objects in process models, e.g., checklists, indicate a reduction of uncertainty. The second metric for this characteristic shows the share of quality documents in relation to all documents in a process.

Individualism refers to the intensity of the interdependence of the members of a cultural group and to the extent of people taking care of each other. In individualistic societies, members look after themselves and their direct environment only, while people in collectivistic societies belong to groups. For example, people working in project teams with a high degree of interdependence and a lot of meetings form a collectivistic group [6]. With regard to the questionnaire, e.g., the item was chosen which asks how important it is for an employee to be surrounded by likeable people in a pleasant working atmosphere (m05). The intensity of interdependence can be identified in a process model by looking at the organizational units of each function. Functions with more than one organizational unit show the interaction among them, e.g., the participants of a meeting. A small share of those functions in relation to all functions indicates a small degree of interdependence and an individualistic attitude. While databases are often used to share information, they also represent a form of individualism if they are not shared among employees. Employees create their own data storage, e.g., a sales person who stores product data on his own device when traveling. The metric measures the share of isolated databases in relation to all databases. A high value indicates that the data is not available to other employees. The degree as to which people take care of each other is not directly measurable, as process models do not show in what way people interact. In each dimension, two metrics were derived on the basis of the definitions and questionnaire items. Thus, our metrics follow the same intention that Hofstede [6] measured with his questionnaire.

3.3 Process of Data Collection

To apply these metrics on process models, their elements have to be analyzed (section 3.1). Figure 2 illustrates an exemplary analysis of a business process model with the cultural values extracted from it. In general, all objects of a process model have to be analyzed either individually or in combination with connected objects. An example is a decision function that can be identified by a subsequent XOR-connector (syntax). Still, the semantics of the function needs to be assessed for final classification to a variable, as also a check function is usually followed by an XOR-connector (see Figure 2). Thus, for each metric, we present a detailed procedure of how to gather the values.

Check functions (CF) are identified by a semantic analysis of two consecutive functions, with the first one producing a result that is checked in the second one. For example, in Figure 2, the quantities and times in the product development list are checked. If this check is performed by an organizational unit of a higher hierarchical level, there will be a value for variable CF_{HL}. Continuing the example, the check of quantities and times is performed by the technical managing director, while, prior to that, the quantities and times were determined by a technician. As a technician is a subordinate of the technical managing director, there is a difference in hierarchy.

Decision functions (DF) are identified by analyzing the syntax and the semantics of the functions and the process flow. Decision functions are followed by alternative process paths indicated by “XOR” or “OR” connectors, e.g., the decision whether a product will be developed or not. Depending on the organizational unit that performs the function, different decision functions can be identified: DF_{HL} indicates a decision made by an organizational unit of a higher hierarchy level than the other organizational units in the process, e.g., a managing director. Quality functions (QF), e.g., the creation of a product specification, can be identified by a semantic analysis of functions. Functions that are performed by only one organizational unit ($F_{OU=1}$) can be identified by a syntactical analysis of the related organizational units that perform a function individually.

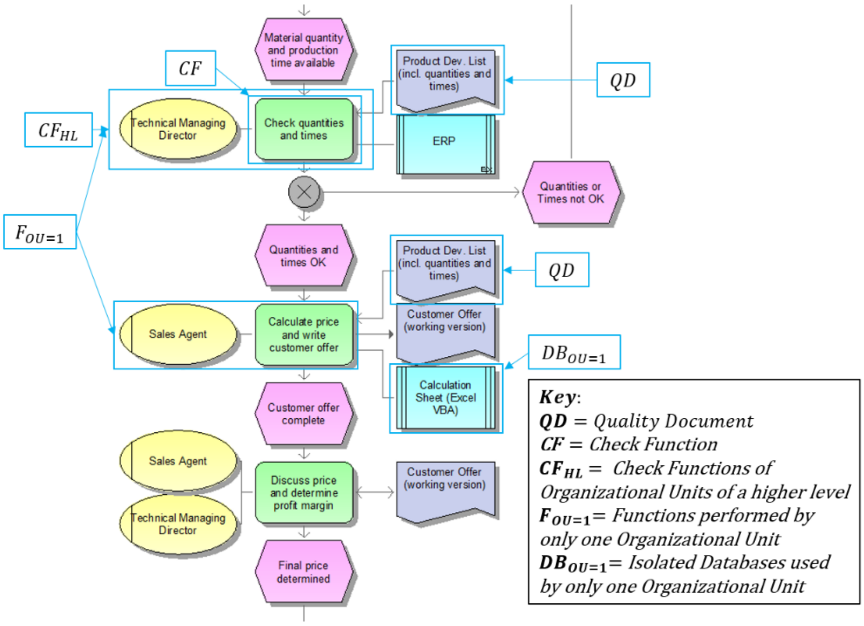


Figure 2. Example of the Data Collection

In addition to functions, information objects are used to measure cultural characteristics. Quality documents (QD) are a subset of all documents (D). They can be identified by selecting those documents that deal with quality issues, e.g., checklists or test reports. Furthermore, isolated databases ($DB_{OU=1}$) can be identified by determining the connection between databases and organizational units. If a database is used by only one organizational unit, it is referred to as an isolated database.

4 Application of Metrics

4.1 Business Processes from Germany and India used for the Application

As cultural values are only meaningful by comparison [8], the metrics derived were applied on business processes of two small and medium-sized enterprises (SMEs) in Germany and India. Both act as suppliers developing and manufacturing similar small electronic components, e.g., transformers, coils, converters.

Out of all business processes, the customer-specific product development process was chosen for the cultural analysis, as both companies develop similar products, which have an analogous complexity and a similar quality required of the output. Due to the high number of products (approx. 350 in each company) that have been developed so far, a high degree of process standardization is available, too, and a similar number of employees performs this process, 11 in the German and 12 in the Indian company. Finally, this process was chosen because it is completely non-automated. Thus, all these similarities ensure the comparability of the German and Indian process so that the cultural characteristics based on human behavior may be fully concentrated on.

The process models we used for our cultural analysis were created in a prior project using the EPC. Even though they were modeled by one and the same modeler, in accordance with identical modeling guidelines and at a comparable level of detail, it was necessary to understand all the details of the processes and check if they were up-to-date. For this purpose, we conducted personal interviews with both the German and the Indian employees involved in the processes. In addition, we were informed about all documents and information systems during on-site visits. On this basis, we slightly revised the models and assured that our analysis was based on the "as-is" models.

The customer-specific product development process starts with a customer asking for a custom-tailored product in the form of a request or tender sent to the SME. This step is followed by an initial design and a subsequent quotation submitted to the customer. In case the customer places an order, the SME designs the product and builds a prototype, which is sent to the customer. After the approval of the prototype, production is planned and initiated. The final steps comprise invoicing and handling of payments. This process is executed in both companies, with some differences regarding the actual process flows and responsibilities. The individual processes are comprised of 60 functions in the German and 71 in the Indian company. Against this background, the values of the variables were independently assessed by two researchers and their results compared afterwards, being identical in almost all cases, showing only slight differences for the rest. The differences were discussed with a third researcher. If no final decision could be made, the differences were discussed with the two companies until a consensus was found to ensure a maximum degree of objectivity when applying the metrics.

4.2 Results of the Application

The values of the variables to calculate the metrics were collected by way of a semantic and syntactical analysis of the processes for both the SMEs as described in section 3.3. We analyzed all objects of both business processes and classified them to the variables that are needed to calculate the metrics. For example, as shown in Figure 2, the Technical Managing Director performs the function *check quantities and times*, which was classified to CF_{HL}. Table 1 shows the defined metrics including a short description, the absolute numbers gained from the process analysis, and the calculated metrics for the German and the Indian company for each cultural dimension. According to our measurement approach, the values of the individual metrics are aggregated for each cultural dimension as well.

Regarding the cultural dimension *Power Distance*, the aggregated value Agg^{PD} is determined by calculating the average of PD_1 and PD_2. Possible and expected values range from 0 (low) to 1 (high Power Distance) for each metric and thus for the aggregated value, too. In the Indian company, 50% of the check functions and even 72% of the decision functions are performed by higher level organizational units, while in the German company only 29% of the checks and 22% of the decisions are made by their counterparts. The aggregated value is 0.61 for the Indian and 0.25 for the German company. Based on these values, we can state that the process reflects a higher Power Distance in the Indian than in the German company.

The aggregation for Uncertainty Avoidance is determined by calculating the average AggUA of the two metrics associated with this cultural dimension. For the two metrics, possible values range from 0 (low) to 1 (high Uncertainty Avoidance), which is also valid for the aggregated value. UA1 reveals that the German process has a share of 28% of quality functions, which means that almost one in three of all functions is quality-related. The Indian process shows a lower share of 17%, which equals less than every fifth function. This tendency is also reflected in the share of quality documents (UA2) with 50% in the German and 16% in the Indian company. In summary, the aggregated values of 0.39 in the German and 0.17 in the Indian company reflect a higher avoidance of uncertainty in the German company.

Table 1. Application of the Derived Metrics (*OU=Organizational Unit)

			Absolute Numbers		Calculated Metric	
	Metric	Description	Ger.	India	Germany	India
Power Distance	$PD_1 = \frac{\lvert CF_{HL} \rvert}{\lvert CF \rvert}$	Check Functions of OUs* of a higher level / Check Functions	$\frac{2}{7}$	$\frac{4}{8}$	0.29	0.50
Power Distance	$PD_2 = \frac{\lvert DF_{HL} \rvert}{\lvert DF \rvert}$	Decision Functions of OUs* of a higher level / Decision Functions	$\frac{2}{9}$	$\frac{8}{11}$	0.22	0.72
Power Distance				Agg^{PD}	0.25	0.61
Uncertainty Avoidance	$UA_1 = \frac{\lvert QF \rvert}{\lvert F \rvert}$	Quality Functions / Functions	$\frac{17}{60}$	$\frac{12}{71}$	0.28	0.17
Uncertainty Avoidance	$UA_2 = \frac{\lvert QD \rvert}{\lvert D \rvert}$	Quality Documents / Documents	$\frac{8}{16}$	$\frac{4}{25}$	0.50	0.16
Uncertainty Avoidance				Agg^{UA}	0.39	0.17
Individualism	$Ind_1 = \frac{\lvert F_{OU=1} \rvert}{\lvert F \rvert}$	Functions performed by only one OU* / Functions	$\frac{50}{60}$	$\frac{50}{71}$	0.83	0.70
Individualism	$Ind_2 = \frac{\lvert DB_{OU=1} \rvert}{\lvert DB \rvert}$	Isolated Databases used by only one OU* / Databases	$\frac{1}{2}$	$\frac{1}{3}$	0.50	0.33
Individualism				Agg^{Ind}	0.67	0.52

The aggregated value of the dimension *Individualism* is determined by building the average (Agg^{Ind}) of the metrics Ind_1 and Ind_2, with 1 indicating a high degree of Individualism and 0 indicating a high degree of Collectivism. The metric values show that in the German process 83% of all functions are performed individually, which indicates a higher degree of individualism than in the Indian process with value of 70% (Ind_1). In the German company, 50% of all databases and 33% in the Indian subsidiary are only used individually (Ind_2). The aggregation Agg^{Ind} results in a value of 0.67 in the German and 0.52 in the Indian company, indicating that the German company is more individualistic than the Indian company.

In summary, all of the derived metrics were applicable on the modeled processes in both the German and the Indian companies, and a clear tendency for each cultural dimensions was identified.

5 Discussion

As we have demonstrated in the previous section, it is possible to identify and measure cultural dimensions in business process models. We operationalized three of Hofstede et al. [8]'s cultural dimensions in process models by deriving metrics analyzing the use and structure of process model elements. Our analysis of the results (see section 4.2) revealed a clear diversity between the German and the Indian process models regarding our cultural metrics. Even though we cannot interpret the derived values as absolute numbers, a clear tendency of the three cultural dimensions for each of the two process models is obvious. For a better interpretation of our results, we opposed them to the findings of the comprehensive study by Hofstede et al. [8] regarding the three cultural dimensions in the two countries Germany and India (see Table 2). Hofstede et al. [8] use a scale from 0 (low) to 100 (high) to rate their cultural dimensions. In comparison, our metrics' values range from 0 (low) to 1 (high). While this means that we cannot directly compare the two scales, we can compare the values of the two countries against their respective scale to reveal a tendency in each cultural dimensions.

Looking at *Power Distance*, Hofstede states a value of 35 for Germany and 77 for India, indicating that Germany has a lower Power Distance than India [6]. Our analysis of the process models comes to a very similar result as our metrics classified the process models of the German company (0.25) showing a lower Power Distance than the ones of the Indian company (0.61). *Uncertainty Avoidance* has a value of 65 in Germany and 40 in India according to Hofstede, displaying a higher avoidance of uncertainty in Germany. These findings are congruent with our measures: The analyzed process models evidence that there is a higher avoidance of uncertainty in the German company (0.39) than the Indian one (0.17). Hofstede's third cultural dimension rates the degree of *Individualism* in a society. His survey states a value of 67 for Germany and 48 for India, reflecting a higher Individualism for Germany in comparison to India. Our results are in accordance with this tendency as we calculated 0.67 for Germany and 0.52 for India using our metrics.

Table 2. Comparison of Calculated Metrics and Survey Results by Hofstede et al. [8]

	Aggregated metrics			Survey by Hofstede et al. [12]		
	Germany		India	Germany		India
Power Distance	0.25	<	0.61	35	<	77
Uncertainty Avoidance	0.39	>	0.17	65	>	40
Individualism	0.67	>	0.52	67	>	48

We see a contribution of our research to both theory and practice. From a scientific point of view, cultural aspects in BPM have mostly been disregarded [18], even though culture has a direct influence on process performance [20]. Our metrics are a new instrument to measure national culture in process models and thus contribute to this research gap of BPM. Up to now, cultural research has largely relied on surveys, which primarily focus on the layers of *Underlying Assumptions* and *Espoused Beliefs and Values*. In contrast, our proposed metrics focus on the layer of *Artifacts*, thus allowing for further aspects of cultural research to be investigated. In addition, due to the fact that our metrics are well-defined and the data collection instructions are quick and easy to apply in a real-life context, the extraction of the underlying process variables is very cost-effective as compared to e.g., questionnaires. Since the necessary information is inherent in BPM systems and data collection is in many cases automated, e.g., gathering the as-is process using process mining algorithms, we therefore expect a broad application basis. Our metrics contribute to practice in other areas, too. Looking at BPM in particular, the metrics can be applied to support the different phases of the BPM life cycle:

(1) In the design phase, the process modeler creates a process that has to be aligned to the national cultures of the users involved. For example, they can be applied in - especially international – mergers and acquisitions (M&As) that play a predominant role in times of globalization [31]. There, cultural “collisions” [10] are observed, which are seen as a possible reason why 50 to 60 percent of M&As fail eventually [31,32]. In this regard, our approach helps to assess the national differences in the companies involved by calculating and comparing the metrics of the processes to be merged. Thus, our metrics help to highlight the cultural compatibility and to detect cultural collisions before the actual process harmonization takes place. A high degree of cultural equality between the companies’ processes will lower integration costs and help employees to easier adapt to the new, joint processes. Finally, the failure rate of M&As can possibly be reduced. The same reasoning is valid for outsourcing initiatives. A close cultural fit of the insourcer’s processes with the ones of the outsourcing company may be seen as an indicator of a seamless integration. Besides the harmonization of two processes, the modelers of a process may calculate our metrics on its own processes and compare the values to benchmarks and reference values. A modeler can then check whether the process they design fits the cultural expectations in advance ensuring a high rate of adoption among the users. In addition, the importance of culturally aligning a business process that involves users from several different countries and cultures is obvious. For example, it is necessary to consider the national culture when designing offshoring processes [10], or when transferring an IT-system to a subsidiary in a foreign country [33].

(2) Our metrics also contribute to the analyze phase of BPM. When a company experiences problems with their process performance, e.g., right after implementing a new process, our metrics help to analyze whether cultural issues are the cause of these shortcomings. In fact, there is empirical evidence, that culture affects the process quality (survey, cf. [20]). Thus, our metrics support a company by indicating whether a business process fits the cultural needs and expectations of all parties involved. For example, when employees from different countries work together, problems due to their different cultural backgrounds may be expected, which affects multinational companies in particular [29]. By comparing the metrics to a benchmark or reference value, possible causes may be identified, which can be further evaluated in an upcoming BPI initiative.

(3) In the improvement phase, the culturally induced causes for a decreasing process performance can be eliminated. In this regard, the descriptions of the metrics help to identify improvement possibilities and, thus, to define measures to align the process to the employees’ expectations. Further, the comparison with benchmarks or reference values give further hints in which cultural dimension the roots of the problems may

particularly be. In summary, as described above, we are convinced that our metrics are a new means to make culture a tangible construct from a BPM perspective.

6 Conclusion

In this paper, we present a measurement approach to identify national culture in business process models. We defined six metrics for the national dimensions *Power Distance, Uncertainty Avoidance* and *Individualism*. Further, these metrics were applied on to business process models of a German and an Indian company. We demonstrated that our metrics are applicable and that they provide further insights into the cultural characteristics of the two companies. We derived tendencies of the characteristics for the German and Indian companies and compared them to those of Hofstede et al. [8]. By applying our metrics in two companies in Germany and India, we can support the theory that it is possible to measure cultural characteristics in business processes.

The proposed approach for measuring national culture based on business process models contributes to both research and practice. So far, cultural research has mainly relied on surveys to measure underlying assumptions and espoused beliefs and values for identifying national culture. We contribute by presenting a measurement approach to measure culture based on documented actions and behavior, thus providing a transparent, independent and unbiased analysis of the underlying cultural dimensions. For practice, we see several areas of contribution. In the design phase of BPM, our metrics support the modelers e.g., by checking whether the process they design fits the cultural expectations of the involved parties. Our metrics can also support in the process analysis phase, e.g., by potentially identifying cultural problems causing performance issues. Furthermore, the analysis of the national culture in process models can be applied when measuring the cultural fit of two companies, e.g., at mergers and acquisitions.

However, our research is not without limitations. The processes in a company are not only subject to national culture, but also to other influencing factors that affect the process design. In each particular setting, these factors need to be analyzed in detail. Especially the interdependence of organizational culture and national culture – even though distinguishable due to different cultural dimensions [17] – needs to be analyzed. Further, as objects in a process model may be seen as boundary objects, literature from this adjacent field will be considered, too, in our upcoming research [34].

In terms of further research, our approach has to be expanded and evaluated: First, for a better interpretation and evaluation of the resulting values, further applications of the metrics in different companies, business processes, and countries are necessary. Second, we will investigate further characteristics which are capable to operationalize characteristics in the dimensions *Power Distance, Uncertainty Avoidance* and *Individualism*. Third, to expand our approach, we will define metrics for further cultural dimensions like e.g., *Masculinity, Indulgence* and *Long-Term Orientation* [8]. Fourth, it is necessary to apply the metrics to further settings to allow for a better evaluation and to establish a benchmark value for each individual country, process type, and industry. Last, even though process models are a means of quality documentation in companies, their quality and topicality cannot always be ensured. Thus, we will define requirements to determine which processes our metrics suit best, e.g., to identify the influence of different modeling notations on our metrics.

References

1. vom Brocke, J., Sinnl, T.: Culture in business process management: a literature review. Business Process Management Journal (BPMJ) 17, 357-378 (2011)
2. Ford, D.P., Connelly, C.E., Meister, D.B.: Information systems research and Hofstede's culture's consequences: an uneasy and incomplete partnership. Engineering Management, IEEE Transactions on 50, 8-25 (2003)
3. Helmreich, R.L.: Anatomy of a system accident: The crash of Avianca Flight 052. The international journal of aviation psychology 4, 265-284 (1994)
4. Brett, J., Behfar, K., Kern, M.C.: Managing multicultural teams. Harvard business review 84, (2006)
5. Schein, E.H.: Organizational Culture and Leadership. John Wiley & Sons, (2004)
6. Hofstede, G.: National Cultures in Four Dimensions: A Research-Based Theory of Cultural Differences among Nations. International Studies of Management and Organization 13, 46-74 (1983)
7. House, R.J., Hanges, P.J., Javidan, M., Dofrman, P.W., Gupta, V.: Culture, Leadership and Organizations - The GLOBE Study of 62 Societies. SAGE Publications, (2004)
8. Hofstede, G., Hofstede, G.J., Minkov, M.: Cultures and Organizations: Software of the Mind. McGraw-Hill, New York, NY, (2010)
9. Braunnagel, D., Johannsen, F., Leist, S.: Coupling and process modeling-An analysis at hand of the eEPC. Paper presented at the Modellierung, Wien, (2014)
10. Weber, Y., Shenkar, O., Raveh, A.: National and corporate cultural fit in mergers/acquisitions: An exploratory study. Management Science 42, 1215-1227 (1996)
11. Grau, C., Moormann, J.: Investigating the Relationship between Process Management and Organizational Culture: Literature Review and Research Agenda. Management and Organizational Studies 1, p1 (2014)
12. Kroeber, A.L., Kluckhohn, C.: Culture: A critical review of concepts and definitions. Peabody Museum of Archaeology & Ethnology, Harvard University 47, (1952)
13. Schmiedel, T., vom Brocke, J., Recker, J.: Culture in Business Process Management: How Cultural Values Determine BPM Success. In: vom Brocke J, Rosemann M (eds) Handbook on Business Process Management 2. Springer, pp 649-664. Heidelberg (2010)
14. Leidner, D.E., Kayworth, T.: A Review of Culture in Information Systems Research: Toward a Theory of Information Technology Culture Conflict. Management Information Systems Quarterly (MISQ) 30, 357-399 (2006)
15. Schneider, S.C.: National vs. corporate culture: Implications for human resource management. Human Resource Management 27, 231-246 (1988)
16. Laurent, A.: The cultural diversity of western conceptions of management. International Studies of Management & Organization 13, 75-96 (1983)
17. Hofstede, G., Neuijen, B., Ohayv, D.D., Sanders, G.: Measuring Organizational Cultures: A Qualitative and Quantitative Study across Twenty Cases. Administrative Science Quarterly 35, 286-316 (1990)

18. Rosemann, M., vom Brocke, J.: The Six Core Elements of Business Process Management. In: vom Brocke J, Rosemann M (eds) Handbook on Business Process Management 1. Springer, pp 105-122. Heidelberg (2010)
19. Niehaves, B., Plattfaut, R., Becker, J.: Business process governance: a comparative study of Germany and Japan. Business Process Management Journal (BPMJ) 18, 347-371 (2012)
20. Grau, C., Moormann, J.: Empirical Evidence for the Impact of Organizational Culture on Process Quality. Paper presented at the European Conference on Information Systems (ECIS), Tel Aviv, Israel, (2014)
21. Schmiedel, T., vom Brocke, J., Recker, J.: Development and Validation of an Instrument to Measure Organizational Cultures' Support of Business Process Management. Information and Management 51, 43-56 (2014)
22. Schmiedel, T., Brocke, J.v., Recker, J.: Which cultural values matter to business process management?: Results from a global Delphi study. Business Process Management Journal (BPMJ) 19, 292-317 (2013)
23. Jayaganesh, M., Shanks, G.G.: A cultural analysis of business process management governanace in Indian organisations. Paper presented at the European Conference on Information Systems (ECIS), Verona, Italy, (2009)
24. Martinsons, M.G., Hempel, P.S.: Chinese Business Process Re-engineering. International Journal of Information Management 18, 393-407 (1998)
25. Griesberger, P., Leist, S., Zellner, G.: Analysis of techniques for business process improvement. Paper presented at the European Conference on Information Systems (ECIS), Helsinki, Finnland, (2011)
26. Myers, M.D., Tan, F.B.: Beyond models of national culture in information systems research. Advanced topics in global information management 2, 14-29 (2003)
27. Thomas, O., Fellmann, M.: Semantic EPC: Enhancing Process Modeling Using Ontology Languages. Paper presented at the Workshop on Semantic Business Process and Product Lifecycle Management (SBPM), (2007)
28. Becker, J., Rosemann, M., Schütte, R.: Grundsätze ordnungsmäßiger modellierung. Wirtschaftsinformatik 37, 435-445 (1995)
29. von Stetten, A., Beimborn, D., Weitzel, T.: Analyzing and Managing the Impact of Cultural Behavior Patterns on Social Capital in Multinational IT Project Teams - A Case Study Approach. Business & Information Systems Engineering (BISE) 43, 137-151 (2012)
30. Harvey, F.: National cultural differences in theory and practice: Evaluating Hofstede's national cultural framework. Information Technology & People 10, 132-146 (1997)
31. Cartwright, S., Cooper, C.L.: The role of culture compatibility in successful organizational marriage. The Academy of Management Executive 7, 57-70 (1993)
32. Weber, Y., Schweiger, D.M.: Top management culture conflict in mergers and acquisitions: A lesson from anthropology. International Journal of Conflict Management 3, 285-302 (1992)
33. Hussain, S.: Technology transfer models across cultures: Brunei-Japan joint ventures. International Journal of Social Economics 25, 1189-1198 (1998)
34. Huvila, I., Dirndorfer, T., Jansen, E.H., McKenzie, P., Westbrook, L., Worrall, A.: Boundary objects in information science research. Proceedings of the American Society for Information Science and Technology 51, 1-4 (2014)

Eine Literaturanalyse zur Integration von Business Rules und Business Process Management

Florian Imgrund[1], Mathäus Malorny[1], Christian Janiesch[1]

[1] Julius-Maximilians-Universität Würzburg, Würzburg, Germany
{florian.imgrund,christian.janiesch}@uni-wuerzburg.de,
mathaeusm@web.de

Abstract. Business Rules (BR) und Business Process Management (BPM) sind eng miteinander verzahnt. Dennoch gibt es nur vergleichsweise wenige Forschungsbeiträge, die sich mit dem Grad der Integration beider Domänen befassen. Auf der Basis des Six-Core-Elements-BPM-Framework wird die Integration von Business Rules mit dem BPM anhand einer State-of-the-Art-Analyse der Literatur untersucht und evaluiert. Die Analyse deckt aktuelle und zukünftig mögliche Entwicklungen auf und stellt im Besonderen einen starken Fokus auf die Erforschung methodischer Ansätze fest. Während das Design und die Implementierung von BR und Geschäftsprozessen im Mittelpunkt stehen, werden organisationale Faktoren wie People und Culture kaum betrachtet. Auch kann ein strategischer Bezug der Integration beider Domänen kaum festgestellt werden.

Keywords: Business Process Management; Business Rules; Literaturanalyse; State-of-the-Art

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Imgrund, F.; Malorny, M.; Janiesch, C. (2017): Eine Literaturanalyse zur Integration von Business Rules und Business Process Management, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 211-225

1 Einleitung

Die informationstechnologischen Entwicklungen der letzten Jahre haben zu einer voranschreitenden Digitalisierung geführt und fordern Unternehmen und ihre Angestellten zunehmend zu einer anhaltenden Effizienzsteigerung, einem erhöhten Automatisierungsgrad sowie einer verstärkten Veränderungsbereitschaft heraus [1]. Business Rules (BR) können an dieser Stelle unterstützen, indem bewährte Vorgehensweisen als Regeln dargestellt und unabhängig von dem ursprünglichen Wissensmedium zentral verfügbar gemacht werden [2]. Sie fungieren als explizite, deklarative Regelwerke [3], die das Geschäft in seiner Struktur und auch in seinem Betrieb anleiten können [4]. Dagegen definiert das Business Process Management (BPM) Verfahrensregeln zur *prozeduralen* Darstellung von Prozessen [5].

Damit die Formulierung der operativen Geschäftsanforderungen via BR eine direkte Auswirkung auf betriebliche Informationssysteme hat, bedarf es einer Verknüpfung beider Domänen [4]. Während der Zusammenhang von Geschäftsregeln und Geschäftsprozessen in [6, 7] vertieft werden kann, gibt es trotz der Abhängigkeiten zwischen BR und BPM in der Literatur derzeit keine systematische Auswertung des aktuellen Status-Quo der Integration beider Domänen.

Auf der Basis einer selektiven Literaturanalyse soll ein thematischer Überblick über aktuelle Forschungsgegenstände gegeben, der State-of-the-Art ihrer Integration herausgearbeitet, aktuelle Entwicklungen aufgezeigt, sowie potentielle Forschungsrichtungen abgeleitet werden. Ausgewählte Literatur wird den sechs Kernelementen des Six-Core-Elements-BPM-Frameworks nach [8] zugeordnet. Jede der sechs Kernkategorien verfügt wiederum über fünf *Capability Areas*, die dessen Inhalte in erhöhter Granularität konsistent wiedergeben. Das Framework bildet die holistische Managementdisziplin des BPM vollumfänglich ab und kann daher als geeignete Bewertungsgrundlage für die Integration von BR und BPM herangezogen werden.

Durch die Zuordnung ist schließlich eine Evaluation des Abdeckungsgrades der BR-Forschung gegenüber der BPM-Forschung möglich. Eine Diskussion, die neben dem Grad der Abdeckung Aussagen bezüglich aktueller und zukünftiger Forschungsschwerpunkte evaluiert, rundet den Beitrag ab.

2 Business Process Management und Business Rules

Das *BPM* findet seine Ursprünge in der statistischen Prozesskontrolle [9], die betriebliche Abläufe in den Mittelpunkt der Aufmerksamkeit des Managements stellt und auf Basis statistischer Methoden Performanceprobleme zu isolieren versucht [10]. In den letzten Jahren ist BPM zu einer umfassenden Managementdisziplin herangewachsen, die einen integrierten Ansatz zur Verbesserung der Unternehmensleistung durch das Management von Geschäftsprozessen bereitstellt [10]. Im Zusammenspiel mit den Grundideen des Business Process Reengineering haben sich die *Identifikation*, *Modellierung*, *Analyse*, *Verbesserung*, *Implementierung*, *Ausführung*, *Monitoring* und *Änderung* von Geschäftsprozessen als Hauptaktivitäten des BPM herausgestellt [10]. BPM stellt ein interdisziplinäres Forschungsfeld dar, das ein breites Spektrum an Aspekten aus unterschiedlichen Forschungsrichtungen zusammenfasst und deshalb in einem ganzheitlichen Kontext betrachtet werden sollte [11–13].

BR sind meist implizit in aufbau- oder ablauforganisatorischen Unternehmensstrukturen enthalten. Sie repräsentieren die Logik des Geschäftsablaufs und bilden das Wissen ab, das in Unternehmensabläufen implementiert ist [2]. BR können als Bestandteil von Geschäftsprozessen betrachtet und aus diesen extrahiert werden [7, 14–16]. BR sind im Gegensatz zur prozeduralen Darstellung von Prozessen innerhalb des BPM deklarativer Natur. Während in der Literatur Interpretationsspielraum bezüglich der Definition von BR herrscht, soll in dieser Arbeit der Ansatz nach Ross [17] vertreten werden. Demzufolge ist eine BR ein Statement, das Aspekte des Geschäfts explizit definiert oder einschränkt. BR sind atomar und können sowohl aus der geschäftsseitigen als auch der informationstechnischen Seite betrachtet werden. Von Halle [18] ergänzt, dass BR einen einschränkenden, leitenden, vorschlagenden, kalkulierenden oder schlussfolgernden Charakter haben können. Prozesse und BR können anhand eines Managementansatzes organisiert werden [19].

BR bilden das Kernwissen über Prozesse ab, während das BPM sequenzielle Abläufe darstellt [2]. D. h., BR stehen in Beziehung zu Wissen und Steuerung, BPM hingegen im Kontext zu Aktivitäten und Ressourcen [7]. Der deklarativen Beschreibung von Tatsachen, Bedingungen und Restriktionen der BR stehen prozedurale Sequenzen von Prozessen gegenüber [6]. BPM ist isoliert betrachtet nicht in der Lage, Entscheidungslogik auf Detailebene abzubilden. Die Integration von BR ermöglicht es im Kontext von BPM dennoch, Regeln in sequenzielle Prozessdiagramme einzubinden und gleichzeitig von einer deklarativen Beschreibung der Geschäftslogik zu profitieren [6, 20, 21].

3 Forschungsmethodik

Um den Integrationsgrad von BR und BPM herauszustellen, wird eine selektive Literaturanalyse durchgeführt. Angereichert mit den Empfehlungen aus [22, 23] baut die Literaturanalyse auf dem in [24] dargestellten Prozess auf, der aus den Phasen *Problemdefinition*, *Literatursuche*, *Literaturauswertung*, *Analyse*, *Interpretation* und *Präsentation* besteht. Um den Anspruch auf Aktualität und Qualität einer wissenschaftlichen Literaturanalyse zu gewährleisten, beschränkt sich die Literatursuche auf Publikationen, die in *Journals*, auf *Konferenzen* oder als *Working Paper* veröffentlicht wurden. Iterative Suchvorgänge zeigten, dass sich die Ergebnisse der Suchanfrage *"Business Rules" AND "*Process*"* bzw. *"Business Rules" AND "Business Process*"* (IEEE Xplore) weitestgehend innerhalb des Betrachtungsgegenstands bewegten. Um ein möglichst umfassendes Bild zu Forschungsanstrengungen im Rahmen des Betrachtungsgegenstandes zu bieten, erfolgte die Suche in acht etablierten Fachdatenbanken (vgl. Tabelle 1) aus den Bereichen der Wirtschaftswissenschaften, Wirtschaftsinformatik und Informatik. Beiträge, deren Publikationsmedium zum Zeitpunkt der Suche nicht Teil der Rankingkategorien A+ bis C des VHB JOURQUAL-Rankings der Teilgebiete Wirtschaftsinformatik [25] oder

Operations Research [26] war, wurden grundsätzlich ausgeschlossen. Ausnahmen wurden für einzelne Publikationen gemacht, die sich entgegen dieses Selektionskriterium dennoch als äußerst relevant herausstellten.

Die Beiträge wurden im Rahmen der Auswertungsphase von den Autoren anhand ihres *Titels*, der *Keywords* und des *Abstracts* unter Zuhilfenahme folgenden Grundsatzes auf Relevanz hin geprüft: Die Domänen BR und BPM oder Aspekte dieser Domänen bilden den zentralen Betrachtungsgegenstand der Forschungsarbeit und stehen im engen oder weiten Zusammenhang miteinander. Ein enger Zusammenhang umfasst die zusammenhängende Betrachtung beider Domänen. Ein weiter Zusammenhäng ergibt sich, wenn eine Domäne den Schwerpunkt der Betrachtung darstellt und die andere Domäne indirekt im Betrachtungsgegenstand enthalten ist.

In der Analysephase ordneten zwei der Autoren die relevanten Titel unter Betrachtung ihrer Kerninhalte unabhängig voneinander einer oder mehreren Kategorien des BPM-Rahmenwerks nach [8] zu. Unstimmigkeiten wurden in einer abschließenden Diskussion gemeinsam konsolidiert. Nicht zuordenbare Titel wurden in eine separate Kategorie „Others" eingeteilt, die nicht Teil des Rahmenwerks ist.

4 Ergebnisse und Auswertung der Literaturanalyse

4.1 Übersicht

Eine Konsolidierung und Bereinigung der 563 Suchtreffer identifiziert insgesamt 78 als relevant eingestufte Forschungsbeiträge, die im Zeitraum von 1994 bis 2016 veröffentlicht wurden. Tabelle 1 zeigt eine Auflistung von Typ und Quelle der Beiträge.

Tabelle 1: Verteilung der Suchtreffer auf Datenbank und Publikationsformat

Datenbank	**[1]**	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**Gesamt**
Suchtreffer:	76	80	24	17	17	23	63	263	**563**
Davon relevant:	6	12	5	2	4	7	13	36	**85 (78*)**
Konferenzbeiträge:	0,0%	1,3%	0,0%	0,0%	0,0%	7,7%	15,4%	46,2%	**70,5%**
Journalartikel	6,4%	9,0%	5,1%	0,0%	0,0%	1,3%	0,0%	0,0%	**21,8%**
Working Paper	0,0%	3,8%	0,0%	1,3%	2,6%	0,0%	0,0%	0,0%	**7,7%**
Konsolidiert:	**6,4%**	**14,1%**	**5,1%**	**1,3%**	**2,6%**	**9,0%**	**15,4%**	**46,2%**	**100,0%**

[1] Business Source Premier, [2] EconBiz [3] ScienceDirect, [4] RePEc, [5] SSRN, [6] AISeL, [7] ACM Digital Library, [8] IEEE Xplore Digital Library **Anzahl relevanter Artikel ohne Duplikate*

Tabelle 2 enthält alle relevanten Forschungsbeiträge. Die Anzahl der *Zitationen* (laut Google Scholar, Stichdatum 18.08.2016) bilden das Ordnungskriterium der Tabelle. Neben der Kategorie *Beitrag* (Referenz auf den Artikel) klärt die Tabelle zudem über das Erscheinungs*jahr* und den *Typ* des Forschungsbeitrags auf: Journal (*J*), Konferenzbeitrag (*K*) oder Working Paper (*WP*). Die Inhalte der *BPM-Kategorie* stützen sich auf die Kategorisierung des angesprochenen Six-Core-Elements-BPM-Frameworks [8] (vgl. Abbildung 1). Die Zuordnung der Beiträge in Tabelle 2 erfolgte mittels Buchstaben (für die Kategorie des BPM-Frameworks) und Zahlen (für die zugehörige Capability Area). Die Kategorie *Others (O)* wie oben eingeführt fasst nicht zuordenbare Forschungsbeiträge zusammen.

Tabelle 2: Auflistung der relevanten Suchtreffer

Beitrag (Referenz)	Jahr	Zitationen	Typ	BPM Kategorie
[27]	2009	297	J	M5
[28]	2006	162	K	M5
[29]	2008	150	J	M3
[6]	2010	146	J	M5
[30]	2004	120	K	M3
[31]	2004	88	WP	I1
[4]	2004	69	J	M5
[32]	2008	63	K	M1
[33]	2007	56	K	M1; M2; I1; I2
[34]	2007	35	WP	M1; M2
[35]	2011	34	J	G5
[36]	2005	30	K	M1
[37]	2011	26	J	M1
[38]	2005	26	J	M1
[39]	2008	25	K	M1
[40]	2010	22	K	M1
[41]	2008	22	K	M5
[42]	2013	20	J	G5
[43]	2011	20	K	M1; M2
[44]	2011	20	K	M1
[7]	2011	18	K	O
[45]	2002	17	K	M5
[46]	2006	14	K	I2
[47]	2012	13	J	M2
[48]	2011	12	J	M1; M2; I1; I2
[49]	2010	12	K	G5
[50]	2011	11	K	G5
[51]	2010	11	K	M2
[52]	2004	11	K	M2
[53]	2013	10	K	M1; M4
[54]	2012	10	K	M1
[15]	2010	9	K	M1
[55]	2010	9	K	M1
[56]	2003	9	J	M1
[57]	2011	8	K	M1
[58]	2009	8	K	M1
[59]	2007	8	K	M2
[60]	2013	7	J	S3
[61]	2011	7	K	S3
[62]	2010	7	K	M1
[63]	2007	7	K	M1
[64]	2007	7	K	M1; M2; M3; M4
[65]	2011	5	K	M1
[66]	2009	5	K	M1
[67]	2011	4	K	M1
[68]	2011	4	K	M5
[69]	2007	4	K	M1
[70]	2014	3	K	M2
[71]	2013	3	K	M1
[72]	2012	3	K	M1
[73]	2010	3	WP	O
[74]	2005	3	K	M1; O
[75]	2014	2	J	M4
[76]	2011	2	J	M1
[77]	2010	2	K	M4
[78]	2010	2	K	M5
[79]	2009	2	K	S2; M1
[80]	2009	2	K	M1
[81]	2005	2	K	M2
[82]	1994	2	K	I1
[83]	2016	1	K	G5
[84]	2015	1	K	G5; M3
[85]	2013	1	WP	G5
[86]	2012	1	K	M3
[87]	2011	1	K	M1
[88]	2010	1	K	I1
[89]	2008	1	K	M1; M2
[90]	2008	1	K	M3
[16]	2015	0	J	M1
[91]	2015	0	K	S1; M4
[92]	2015	0	J	M3
[93]	2013	0	WP	P2
[94]	2011	0	K	M1
[95]	2011	0	K	M1
[96]	2010	0	K	S3
[97]	2010	0	J	M1
[98]	2008	0	WP	O
[99]	2008	0	K	S3

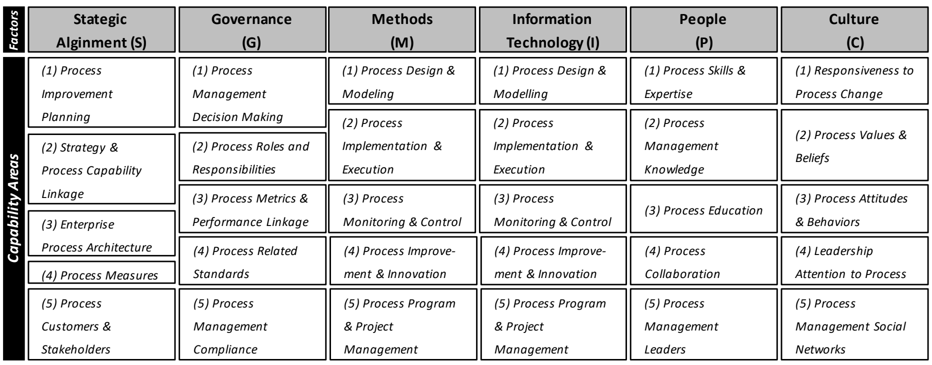


Abbildung 1: The Six Core Elements of BPM nach [8]

Beiträge auf Konferenzen sind derzeit das wichtigste Medium (70,5 %). Etwa jede fünfte Publikation (21,8 %) wird in einem Journal veröffentlicht. Die übrigen Publikationen sind Working Paper.

Weiter fällt eine ungleichmäßige Verteilung der Forschungsbeiträge in die Kategorien des BPM-Frameworks auf. Während viele Autoren im Bereich der *Methods (M)* publizieren, wird den Kategorien *People (P)* und *Culture (C)* kaum bzw. keine Beachtung geschenkt. Die Anzahl sowie die Zitationszahlen der identifizierten Forschungsbeiträge zeigen, dass BR im Kontext des BPM bereits seit 2004 diskutiert werden. Eine vorangeschrittene Integration kann jedoch lediglich in einzelnen Subkategorien (vgl. Kapitel 4.2) erkannt werden. Im zeitlichen Verlauf lässt sich ein erster Forschungsschwerpunkt im Jahr 2011 erkennen, der jedoch nicht weiter anhält.

4.2 Darstellung des State-of-the-Art

Der State-of-the-Art der Integration von BR und BPM wird in Abbildung 2 Kategorie-basiert dargestellt. Aufgrund von Mehrfachzuordnungen übersteigt die Gesamtanzahl der in Abbildung 2 aufgeführten Beiträge (95) die tatsächliche Menge der Beiträge (78).

Die visualisierte Darstellung zeigt, dass sich ein Großteil der Forschung auf die Erforschung und Evaluation methodischer Ansätze *(Methods)* fokussiert (72,63 %), während die übrigen Kernkategorien *Strategic Alignment*, *Governance*, *Information Technology*, *People* und *Culture* mit einer Zuordnungsrate von insgesamt 23,16 % vergleichsweise geringe Beachtung finden. Beiträge, die keine Zuordnung finden konnten, sind in der Kategorie *Others* (4,21 %) eingeordnet.

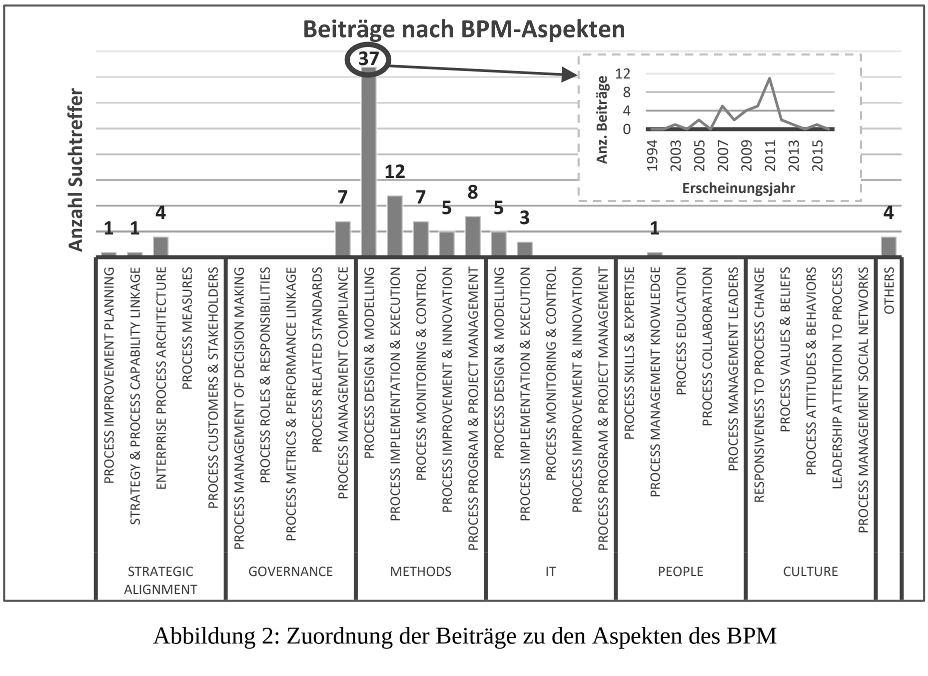


Abbildung 2: Zuordnung der Beiträge zu den Aspekten des BPM

Strategic Alignment

Die Kategorie *Strategic Alignment* beschäftigt sich mit der Ausrichtung von Unternehmensabläufen an die Unternehmensstrategie.

Im Kontext zu BR stehen besonders *integrierte Architekturmodelle* im Fokus, die durch die Integration unterschiedlicher Modellierungsansätze für Geschäftsprozesse und BR [99] sowie die Verbesserung der Adaptionsfähigkeit von Unternehmen anhand von Konzepten für serviceorientierte Architekturen [35, 60, 96] motiviert werden. Integrierte Architekturmodelle sollen einer mangelhaften Interoperabilität, langen Umsetzungszeiten und hohen Kosten entgegenwirken [60] und somit den Grundstein für eine bessere vertikale und horizontale Integration von Strategievorhaben legen. Die Integration von BR vereinfacht die Implementierung neuer Abläufe, reduziert Kosten und verbessert das Alignment zwischen Strategie und Geschäftsprozessen [60].

Governance

Aktivitäten im Bereich *Governance* beschäftigen sich vorwiegend mit der Festlegung von Rollen und Verantwortlichkeiten sowie dem Prozess der Entscheidungsfindung innerhalb des BPM [8]. BR thematisieren in diesem Kontext die Unterkategorie *Process Management Compliance*, indem sie ein regelbasiertes Monitoring sowie eine regelbasierte Validierung von Geschäftsprozessen ermöglichen [35, 42, 49, 84].

Compliance Prüfungen, die sich den Mitteln des Process Mining bedienen, können so um regelbasierte Ansätze erweitert und die Auswertung von Event-Logs durch den Einsatz von BR beschleunigt werden [42]. Zudem können BR zur einfachen und schnellen Validierung von Prozessabläufen im Rahmen von Modellierungstätigkeiten herangezogen werden [49]. Außerhalb des Compliance Checking finden BR in dieser Kategorie bisher keine weitere Aufmerksamkeit. Aufgrund des deklarativen Charakters von BR könnten weiterführende Diskussionen besonders in der Capability Area *Process Management of Decision Making* Mehrwert schaffen.

Methods

Als *Methoden* des BPM werden Techniken und Instrumente beschrieben, die eine Ausführung spezifischer Aktivitäten entlang der Lifecycle-Phasen des BPM erlauben [8]. In dieser Kategorie teilen sich 69 Artikel auf jede der fünf Capability Areas auf.

Besondere Aufmerksamkeit gilt dabei dem *Design & Modeling*, in dessen Zusammenhang Fragestellungen zur Darstellung von BR innerhalb des BPM diskutiert werden [33, 36, 39]. In diesem Kontext stellt die Object Management Group mit der Decision Model and Notation [100] einen Standard zur Modellierung von Entscheidungsszenarien bereit. Abbildung 2 lässt erkennen, dass die Diskussion in dieser Capability Area ab dem Jahr 2011 abflacht. Zur Mühlen und Indulska [6] kommen zu der Feststellung, dass eine Kombination von SRML und BPMN für eine integrierte Modellierung von BR und Geschäftsprozessen als adäquat betrachtet werden kann. Dies darf jedoch im Hinblick auf diese Subkategorie nicht mit einer abgeschlossenen Erforschung gleichgesetzt werden.

Themenstellungen, die sich ergänzend mit der Ausführbarkeit von BR im Kontext des BPM beschäftigen, finden sich in der Capability Area *Implementation & Execution* wieder [34, 43, 47, 51, 52, 101]. Hier geht es im Besonderen um die Fähigkeit von BR, ausführbare Geschäftsprozesse zu implementieren.

Veröffentlichungen der Unterkategorie *Process Monitoring & Control* beschäftigen sich vornehmlich mit der Verifikation von Geschäftsprozessen auf Basis semantischer Rahmenbedingungen, die in Form von BR repräsentiert werden [29, 30, 84]. Im Bereich *Process Improvement & Innovation* geht es um die Adaption und Verbesserung von Geschäftsprozessen unter Zuhilfenahme von BR [53, 75, 77]. Zudem werden Möglichkeiten des Outsourcings von Geschäftsprozessen in die Cloud mithilfe sogenannter Business Rule Management Systeme diskutiert [91].

Beiträge der Capability Area *Program & Project Management* untersuchen verschiedene Modellierungsansätze zu BR und BPM und stellen diese unter Berücksichtigung einer anschließenden Bewertung einander gegenüber [6, 27, 28, 45].

Information Technology

Die *Information Technology* befasst sich mit Informationssystemen sowie Software und Hardware, die eine Durchführbarkeit der Methoden des BPM gewährleisten [8].

In Bezug auf BR werden Programme zur Adaption von Geschäftsprozessen [46] sowie Softwarelösungen für das Management von BR evaluiert [88]. Zudem wird die toolbasierte Modellierung von wissensintensiven Unternehmensabläufen diskutiert [31, 82] und Methoden sowie Softwaretools zur Transformation von BR [51] bewertet. Damit werden insbesondere die Capability Areas *Design & Modelling* und *Implementation & Execution* bedient. Eine Diskussion von BR im Bereich *Monitoring & Control* sowie dem eher organisationalen *Program & Project Management* findet nicht statt.

People

Die Kategorie *People* betrachtet Menschen als wesentlichen Bestandteil in Unternehmensabläufen und beschäftig sich mit deren Fähigkeiten und Erfahrungen in Bezug auf spezifische Prozessanforderungen [8]. Im Kontext zu BR kann ein thematischer Bezug lediglich zu einer Fallstudie hergestellt werden, die Herausforderungen der Implementierung von BR-Ansätzen u. a. vor dem Hintergrund der Nutzerakzeptanz sowie der Kosten-Nutzen-Effizienz untersucht [93]. Nichtsdestotrotz wäre insbesondere die Abbildung von *Skills & Expertise* ein Aspekt der sich gut zur Ressourcensteuerung von BPM mit BR unterstützen ließe.

Culture

Die Haltung gegenüber Prozessverbesserungsmaßnahmen sowie die Werte eines Unternehmens werden im BPM-Framework unter der Kategorie *Culture* diskutiert. In der durchgeführten Literaturanalyse konnten jedoch keine Zusammenhänge zwischen BR und Inhalten dieser Kernkategorie gefunden werden.

Others

Arbeiten, die sich den Aspekten im Framework nicht zuordnen ließen, wurden in die Kategorie *Others* eingeordnet. Die Kategorie Others ist kein Bestandteil des eigentlichen Frameworks. Hier wird beispielsweise die Integration von BR und BPM betrachtet, indem eine Kategorisierung von BR ausgearbeitet wird [7].

4.3 Diskussion

Die vorangegange Literaturanalyse zeigt, dass BR aufgrund ihrer deskriptiven Natur als ein adäquates Werkzeug zur Darstellung, Umsetzung und Überwachung von strategischen Initiativen fungieren. BR können Prozessarchitekturen durch die Integration ihrer eingekapselten Logik agiler gestalten und durch ihre steuernde Charakteristik bei der Ausführung von Prozessen unterstützen. Ursache-Wirkungs-Analysen profitieren von der inhärenten Logik der BR, während ihr lenkender und einschränkender Charakter in der Entscheidungsfindung bei der Ausgestaltung von Prozessen unterstützen kann. Auch Prozessmanagement-Initiativen profitieren von BR, indem diese in der Entscheidungsmodellierung Berücksichtigung finden. Entsprechend findet eine integrierte Betrachtung von BR und BPM vorwiegend im Bereich Methods statt. Eine isolierte Betrachtung der Domänen BPM und BR würde höchstwahrscheinlich zu einem identischen Ergebnis kommen.

Neben dieser stark funktional orientierten Betrachtung, die sich in den meisten Anwendungsfällen auf die Wertschöpfung des Unternehmens beschränkt, werden BR im Kontext organisationaler Faktoren (*People* und *Culture*) kaum diskutiert. Da BR für Menschen lesbare Regeln abbilden und von Fachabteilungen gepflegt werden sollen, sind diese beiden Kategorien jedoch hochrelevant. Die Autoren empfinden es daher als nötig, dass sich neben der methodischen Auseinandersetzung auch eine Diskussion bzgl. organisationaler Aspekte von BR etabliert. In einer isolierten Betrachtung zeigt das BPM im Gegensatz zu BR bereits eine erhöhte Sensibilität für entsprechende Faktoren. So stellen diverse Publikationen die Berücksichtigung der Faktoren *People* und *Culture* als erfolgskritisch hinsichtlicher der Erreichung BPM-spezifischer Ziele heraus [102–106]. Analog zum BPM müssen auch BR-spezifische Methodiken zunächst von ihren Anwendern (People) interpretiert werden, bevor eine Umsetzung in einer organisationalen Umgebung (Culture) erfolgen kann.

Die Literaturanalyse deutet jedoch darauf hin, dass BR und damit auch die Integration von BR und BPM den Schritt von einem technisch getriebenen Thema hin zu einer organisational verankerten Methodologie noch nicht vollzogen hat. Diese These stützt auch der Fakt, dass sich BR im Bereich *Governance* bislang nur wenig etabliert haben, da es hier in aller Regel um nicht-automatisierbare Entscheidungen des BPM oder BR-Management geht.

Wir schlagen daher eine Erweiterung des derzeit eher methodisch orientierten Forschungsdialogs auf den Aufbau eines adäquaten Verständnises zur ganzheitlichen, integrierten Betrachtung von BR und BPM vor, um somit die organisationale Einordnung der Themen in den Unternehmenskontext zu fördern.

5 Zusammenfassende Erläuterungen

Die vorliegende Arbeit gibt einen thematischen Überblick zur Integration von BR und BPM und stellt aktuelle Entwicklungen sowie künftig mögliche Forschungsrichtungen heraus. Die Auswertungen deuten auf eine vorangeschrittene, aber ungleichmäßig verteilte Diskussion der Integration von BR und BPM hin. Ein großer Fokus liegt auf der Erforschung von Methoden, die insbesondere spezifische Darstellungs- und Modellierungsformen von BR im Kontext des BPM betrachten.

Die Auswertung der Literaturanalyse zeigt deutlich, dass die Integration von BR und BPM keinesfalls als abgeschlossen angesehen werden kann und weiterer Forschungsbedarf besonders in geringer erforschten Kerngebieten wie *People* oder *Culture* wünschenswert wäre. Darüber hinaus sollten vorhandene Forschungsthemen validiert und erweitert werden. Die Literaturanalyse zeigt zudem Raum für eine Vertiefung von Ansätzen bzgl. der Identifikation und Extraktion von BR sowie dessen Einsatz zur Identifikation von Prozessschwachstellen und im Bereich der Governance.

References

1. Halle, B. von: Business Rules Applied. Business better Systems Using the Business Rules Approach. Wiley Computer Publishing, New York, NY (2002)
2. Ross, R.G.: Principles of the Business Rule Approach. Addison-Wesley, Boston, MA (2003)
3. Ronald G. Ross: The Business Rules Manifesto, http://www.businessrulesgroup.org/brmanifesto.htm (Abgerufen am: 21. Juli 2016)
4. Kovacic, A.: Business Renovation: Business Rules (Still) the Missing Link. Business Process Management Journal 10, 158–170 (2004)
5. van der Aalst, Wil M. P.: Business Process Management. A Comprehensive Survey. ISRN Software Engineering 2013, 1–37 (2013)
6. Zur Muehlen, M., Indulska, M.: Modeling Languages for Business Processes and Business Rules. A Representational Analysis. Information Systems 35, 379–390 (2010)
7. Zoet, M., Versendaal, J., Ravesteyn, P., Welke, R.: Alignment of Business Process Management and Business Rules. In: Proceedings of the 19th European Conference on Information Systems (ECIS), pp. 1 - 12. Helsinki (2011)
8. Rosemann, M., Vom Brocke, J.: The Six Core Elements of Business Process Management. In: Vom Brocke, J., Rosemann, M. (eds.) Handbook on Business Process Management 1, pp. 107–122. Springer, Berlin (2010)
9. Shewhart, W.A., Deming, W.E.: Statistical Method from the Viewpoint of Quality Control. Dover Books on Mathematics, New York, NY (1986)
10. Hammer, M.: What is Business Process Management? In: Vom Brocke, J., Rosemann, M. (eds.) Handbook on Business Process Management 1, pp. 3–16. Springer, Berlin (2010)
11. Margherita, A.: Business Process Management System and Activities. Business Process Management Journal 20, 642–662 (2014)
12. Kohlborn, T., Vom Brocke, J., Schmiedel, T., Recker, J., Trkman, P., Mertens, W., Viaene, S., Mueller, O.: Ten Principles of good Business Process Management. Business Process Management Journal 20, 530–548 (2014)

13. Moore, C., Saxena, R., Lee, D., Powell, E.: BPM CBOK. Version 3.0. Association of Business Process Management International, Lexington, KY (2013)
14. Tang, W., Qin, J.: The Research of a Business Rule Management Method for Virtual Organization. In: Proceedings of the 5th International Conference on MEMS NANO, and Smart Systems (ICMENS), pp. 25–28. Dubai (2009)
15. Levina, O., Holschke, O., Rake-Revelant, J.: Extracting Business Logic from Business Process Models. In: Proceedings of the 2nd International Conference on Information Management and Engineering (ICIME), pp. 289–293. Chengdu (2010)
16. Polpinij, J., Ghose, A., Dam, H.K.: Mining Business Rules from Business Process Model Repositories. Business Process Management Journal 21, 820–836 (2015)
17. Business Rules Group: Defining Business Rules: What Are They Really? Final Report. Revision 1.3, http://www.businessrulesgroup.org/first_paper/br01c0.htm (Abgerufen am: 11.03.2016)
18. Halle, B. von: The Business Rule Revolution. Running Business the Right Way. Happy About, Cupertino, CA (2006)
19. Kardasis, P., Loucopoulos, P.: Expressing and Organising Business Rules. Information and Software Technology 46, 701–718 (2004)
20. Schlosser, S., Baghi, E., Otto, B., Oesterle, H.: Toward a Functional Reference Model for Business Rules Management. In: Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS), pp. 3837–3846. Waikoloa, HI (2014)
21. Nelson, M.L., Peterson, J., Rariden, R.L., Sen, R.: Transitioning to a Business Rule Management Service Model. Case studies from the Property and Casualty Insurance Industry. Information & Management 47, 30–41 (2010)
22. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26, xiii–xxiii (2002)
23. Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., Cleven, A.: Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. Communications of the Association for Information Systems (CAIS) 37, 205–224 (2015)
24. Valentine, J.C., Hedges, L.V., Cooper, H.M.: The Handbook of Research Synthesis and Meta-Analysis. Russell Sage Foundation, New York, NY (2009)
25. Verband der Hochschullehrer für Betriebswirtschaft e.V.: VHB-JOURQUAL 3. Teilrating Wirtschaftsinformatik, http://vhbonline.org/service/jourqual/vhb-jourqual-3/teilrating-wi (Abgerufen am: 06.04.2016)
26. Verband der Hochschullehrer für Betriebswirtschaft e.V.: VHB-Jourqual 3. Teilrating Operations Research, http://vhbonline.org/vhb4you/jourqual/vhb-jourqual-3/teilrating-or/ (Abgerufen am: 22. Juli 2016)
27. Green, P., Indulska, M., Recker, J., Rosemann, M.: Business Process Modeling- A Comparative Analysis. Journal of the Association for Information Systems 10, 333–363 (2009)
28. Green, P., Indulska, M., Recker, J., Rosemann, M.: How Good is BPMN Really? Insights from Theory and Practice. In: Proceedings of the 14th European Conference on Information Systems (ECIS), 135, pp. 1–12. Gothenburg (2006)
29. Ly, L.T., Rinderle, S., Dadam, P.: Integration and Verification of Semantic Constraints in Adaptive Process Management Systems. Data & Knowledge Engineering 64, 3–23 (2008)
30. Lazovik, A., Aiello, M., Papazoglou, M.: Associating Assertions with Business Processes and Monitoring their Execution. In: Proceedings of the 2nd International Conference on Service Oriented Computing (ICSOC), pp. 94–104. New York, NY (2004)

31. Gronau, N., Weber, E.: Management of Knowledge Intensive Business Processes. In: Desel, J., Pernici, B., Weske, M. (eds.) Business Process Management, 3080, pp. 163–178. Springer, Berlin (2004)
32. van Eijndhoven, T., Iacob, M., Ponisio, M.L.: Achieving Business Process Flexibility with Business Rules. In: Proceedings of the 12th EDOC Conference, pp. 95–104. Munich (2008)
33. Graml, T., Bracht, R., Spies, M.: Patterns of Business Rules to Enable Agile Business Processes. In: Proceedings of the 11th EDOC Conference, p. 365. Annapolis, MD (2007)
34. Goedertier, S., Haesen, R., Vanthienen, J.: EM-BrA2CE v0.1. A Vocabulary and Execution Model for Declarative Business Process Modeling. SSRN Electronic Journal (2007)
35. Weigand, H., van den Heuvel, W., Hiel, M.: Business Policy Compliance in Service-oriented Systems. Information Systems 36, 791–807 (2011)
36. Colomb, R., Governatori, G., Padmanabhan, V., Rotolo, A., Sadiq, S.: Process Modelling: The Deontic Way. In: Proceedings of the 3th APCCM, pp. 75–84. Darlinghurse (2006)
37. Döhring, M., Zimmermann, B.: vBPMN: Event-Aware Workflow Variants by Weaving BPMN2 and Business Rules. In: Halpin, T.A. (ed.) Enterprise, Business-process and Information Systems Modeling, 81, pp. 332–341. Springer, Berlin (2011)
38. Kardasis, P., Loucopoulos, P.: A Roadmap for the Elicitation of Business Rules in Information Systems Projects. Business Process Management Journal 11, 316–348 (2005)
39. Zur Muehlen, M., Indulska, M., Kittel, K.: Towards Integrated Modeling of Business Processes and Business Rules. In: Proceedings of the 19th Australasian Conference on Information Systems (ACIS), pp. 689–699. Christchurch (2008)
40. Kalsing, A.C., do Nascimento, G.S., Iochpe, C., Thom, L.H.: An Incremental Process Mining Approach to Extract Knowledge from Legacy Systems. In: Proceedings of the 14th EDOC Conference, pp. 79–88. Vitoria (2010)
41. Nelson, M.L., Rariden, R.L., Sen, R.: A Lifecycle Approach towards Business Rules Management. In: Proceedings of the 41th Hawaii International Conference on System Sciences (HICSS), p. 113. Waikoloa, HI (2008)
42. Caron, F., Vanthienen, J., Baesens, B.: Comprehensive Rule-based Compliance Checking and Risk Management with Process Mining. DSS 54, 1357–1369 (2013)
43. Di Bona, D., Lo Re, G., Aiello, G., Tamburo, A., Alessi, M.: A Methodology for Graphical Modeling of Business Rules. In: European Modelling Symposium, pp. 102–106. Madrid (2011)
44. Milanovic, M., Gasevic, D., Rocha, L.: Modeling Flexible Business Processes with Business Rule Patterns. In: Proceedings of the 15th EDOC Conference, pp. 65–74. Helsinki (2011)
45. Groznik, A., Kovacic, A.: Business Renovation: From Business Process Modelling to Information System Modelling. In: Proceedings of the 24th ITI, pp. 405–409. Cavtat (2002)
46. Rosenberg, F., Nagl, C., Dustdar, S.: Applying Distributed Business Rules - The VIDRE Approach. In: Proceedings of the SCC, pp. 471–478. Chicago, IL (2006)
47. Malik, S., Bajwa, I.S.: Back to Origin: Transformation of Business Process Models to Business Rules. In: Rosa, M., Soffer, P. (eds.) Business Process Management Workshops, 132, pp. 611–622. Springer, Berlin (2013)
48. Nalepa, G.J., Kluza, K., Ernst, S.: Modeling and Analysis of Business Processes with Business Rules. In: Beckmann, J.A. (ed.) Business Process Modeling: Software Engineering, Analysis and Applications, pp. 135–156. Nova Science Publishers, Hauppauge, NY (2011)

49. Moura Araujo, B. de, Schmitz, E.A., Correa, A.L., Alencar, A.J.: A Method for Validating the Compliance of Business Processes to Business Rules. In: Proceedings of the ACM Symposium on AC, p. 145. Sierre (2010)
50. Becker, J., Bergener, P., Delfmann, P., Weiss, B.: Modeling and Checking Business Process Compliance Rules in the Financial Sector. In: Proceedings of the 32th International Conference on Information Systems (ICIS), 1–19. Shanghai (2011)
51. Steen, B., Pires, L.F., Iacob, M.: Automatic Generation of Optimal Business Processes from Business Rules. In: Proceedings of the 11th EDOC Conference, pp. 117–126. Vitoria (2010)
52. Groznik, A., Kovacic, A.: The Business Rule-Transformation Approach. In: Proceedings of the 26th International Conference on Information Technology Interfaces (ITI), 1, pp. 113–117. Cavtat (2004)
53. Huang, Y., Feng, Z., He, K.: Ontology-Based Configuration for Service-Based Business Process Model. In: Proceedings of the SCC, pp. 296–303. Santa Clara, CA (2013)
54. Ligěza, A., Kluza, K., Potempa, T.: AI Approach to Formal Analysis of BPMN Models. Towards a Logical Model for BPMN Diagrams. In: Proceedings of the FedCSIS, pp. 931–934. Wroclaw (2012)
55. Polpinij, J., Ghose, A.K., Dam, H.K.: Business Rules Discovery from Process Design Repositories. In: Proceedings of the 6th World Congress on Services, pp. 614–620. Miami, FL (2010)
56. McDermid, D.C.: Integrated Business Process Management: Using State-Based Business Rules to Communicate between Disparate Stakeholders. In: van der Aalst, W. (ed.) Business Process Management, 2678, pp. 58–71. Springer, Berlin (2003)
57. Vasilecas, O., Normantas, K.: Deriving Business Rules from the Models of Existing Information Systems. In: Proceedings of the 12th CompSysTech, pp. 95–100. Vienna (2011)
58. Nalepa, G.J., Mach, M.A.: Business Rules Design Method for Business Process Management. In: Proceedings of the International Multiconference on Computer Science and Information Technology (IMCSIT), pp. 165–170. Mragowo (2009)
59. Milanovic, M., Kaviani, N., Gasevic, D., Giurca, A., Wagner, G., Devedzic, V., Hatala, M.: Business Process Integration by Using General Rule Markup Language. In: Proceedings of the 11th EDOC Conference, p. 353. Annapolis, MD (2007)
60. Gong, Y., Janssen, M.: An Interoperable Architecture and Principles for Implementing Strategy and Policy in Operational Processes. Computers in Industry 64, 912–924 (2013)
61. Gong, Y., Janssen, M.: Creating Dynamic Business Processes using Semantic Web Services and Business Rules. In: Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance (ICEGOV), p. 249. Tallinn (2011)
62. Zhao, K., Ying, S., Zhang, L., Hu, L.: Achieving Business Process and Business Rules Integration using SPL. In: Proceedings of the FITME, pp. 329–332. Changzhou (2010)
63. Thanawut A., Wiwat V.: A Development of Business Rules with Decision Tables for Business Processes. In: Proceedings of the TENCON, pp. 1–4. Taipei (2007)
64. Ali, S., Torabi, T., Soh, B.: Rule Component Specification for Business Process Deployment. In: Proceedings of the 18th International Conference on Database and Expert Systems Applications (DEXA), pp. 595–599. Regensburg (2007)
65. Thi, T.T.P., Helfert, M., Hossain, F., Le Dinh, T.: Discovering Business Rules from Business Process Models. In: Proceedings of the 12th CompSysTech, p. 259. Vienna (2011)
66. Milanović, M., Gašević, D., Wagner, G., Devedžić, V.: Modeling Service Orchestrations with a Rule-enhanced Business Process Language. In: Conference of the Center for Advanced Studies, p. 70. Ontario (2009)

67. Wu, Z., Yao, S., He, G., Xue, G.: Rules Oriented Business Process Modeling. In: Proceedings of the International Conference on Internet Technology and Applications, 2011, pp. 1–4. Wuhan (2011)
68. Koehler, J.: The Process-Rule Continuum - Can BPMN & SBVR Cope with the Challenge? In: Proceedings of the 13th CEC, pp. 302–309. Luxembourg-Kirchberg (2011)
69. Corradini, F., Meschini, G., Polzonetti, A., Riganelli, O.: A Rule-Driven Business Process Design. In: Proceedings of the 29th ITI, pp. 401–406. Cavtat (2007)
70. Bonais, M., Nguyen, K., Pardede, E., and Rahayu, W.: A Formalized Transformation Process for Generating Desing Models from Business Rules. In: Proceedings of te 18th Pacific Asia Conference on Information Systems (PACIS), 21–35. Chengdu (2014)
71. Schrefl, M., Neumayr, B., Stumptner, M.: The Decision-scope Approach to Specialization of Business Rules: Application in Business Process Modeling and Data Warehousing. In: Proceedings of the 9th APCCM, pp. 3–18. Adelaide (2013)
72. Sriganesh, S., Ramanathan, C.: Externalizing Business Rules from Business Processes for Model based Testing. In: Proceedings of the ICIT, pp. 312–318. Athen (2012)
73. Bollen, P.: BPMN: A Meta Model for the Happy Path. Maastricht (2010)
74. Ali, S., Soh, B., Torabi, T.: An Agent Oriented Framework for Automating Fules in Business Processes. In: Proceedings of the 3rd INDIN, pp. 252–256. Perth (2005)
75. Malihi, E., Aghdasi, M.: A Decision Framework for Optimisation of Business Processes Aligned with Business Goals. International Journal of Business Information Systems 15, 22–42 (2014)
76. Sellner, A., Zinser, E.: Establishing Conceptual and Functional Links between S-BPM and Business Rules. In: Fleischmann, A., Schmidt, W., Stary, C., Obermeier, S., Brger, E. (ed.) Subject-oriented Business Process Management, 138, pp. 121–133. Springer, Berlin (2011)
77. Aghdasi, M., Malihi, S.E.: Rule Based Business Process Optimization. In: Proceedings of the IEEM, pp. 305–309. Macao (2010)
78. Levina, O., Kubicki, A., Holschke, O., Rake-Revelant, J.: Exploring the Impact of Business Rules Consideration on Processes in E-Business. In: Proceedings of the 2nd ICIME, pp. 275–279. Chengdu (2010)
79. Boukhebouze, M., Amghar, Y., Benharkat, A., Maamar, Z.: Towards an Approach for Estimating Impact of Changes on Business Processes. In: Proceedings of the 12th CEC, pp. 415–422. Vienna (2009)
80. Yu, J., Sheng, Q.Z., Falcarin, P., Morisio, M.: Weaving Business Processes and Rules: A Petri Net Approach. In: Yang, J. (ed.) Information Systems, 20, pp. 121–126. Springer, Heidelberg (2009)
81. Ali, S., Soh, B., Torabi, T.: Design and Implementation of a Model for Business Rules Automation. In: Proceedings of the 3rd ICITA, pp. 40–45. Sydney (2005)
82. Yigang C., Yutaka U.: EUOBFE-An End User Oriented CASE Tool for Business Process Reengineering. In: Proceedings of the 18th COMPSAC, pp. 346–351. Taipei (1994)
83. Pham, T.A., Le Thanh, N.: An Ontology-based Approach for Business Process Compliance Checking. In: Proceedings of the 10th ICUIMC, pp. 1–6. Danang (2016)
84. Kotamarthi, K., Wang, X., Grossmann, G., Sheng, Q.Z., Indrakanti, S.: A Framework Towards Model Driven Business Process Compliance and Monitoring. In: Proceedings of the 19th EDOCW, pp. 24–32. Adelaide (2015)
85. Caron, F., Vanthienen, J., Baesens, B.: Advances in Rule-Based Process Mining. Applications for Enterprise Risk Management and Auditing. SSRN Electronic Journal (2013)
86. Thirumaran, M., Dhavachelvan, P., Aishwarya, D.: Business Workflow Growth Rate Analysis using Cellular Automata. In: Proceedings of the 2nd CCSEIT, pp. 315–321. Coimbatore (2012)

87. Xin, Y., Chao, Z., Xinghua, B., Yanxin, C.: An Inter-organizational Dynamic Business Process Oriented Decision-Making Behavioral Rule Description Method. In: International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), pp. 90–93. Lyon (2011)
88. Koshkina, M., Huynh, K., Zhao, Y.: Achieving Business Agility with WebSphere ILOG JRules and WebSphere BPM. In: Proceedings of the CAS, p. 360. Toronto (2010)
89. Li Zhao, Li, F.: Statistical Machine Learning in Natural Language Understanding: Object Constraint Language Translator for Business Process. In: KAM Workshop, pp. 1056–1059. Wuhan (2008)
90. da Silveira, D.S., Netto, P.O.B., Schmitz, E.A.: Um Método de Validação da Conformidade Entre Processos e Regras de Negócio Através da Animação. (In Portuguese). In: Companion Proceedings of the XIV Brazilian Symposium on Multimedia and the Web, pp. 279–286. Vila Velha (2008)
91. van Grondelle, J., Liefers, R., Versendaal, J.: Cloud Oriented Business Process Outsourcing using Business Rule Management. In: BLED, 6. Bled (2015)
92. Lam, V.: Constraint-based Reasoning on Declarative Process Execution with the Logics Workbench. Business Process Management Journal 21, 586–609 (2015)
93. Stohr, E.A., Huang, W.: Business Rules Management: Implementation and Evaluation. SSRN Electronic Journal (2012)
94. A.R. Goncalves, J.C. de, Santoro, F.M., Baiao, F.A.: Collaborative Narratives for Business Rule Elicitation. In: Proceedings of the SMC, pp. 1926–1931. Anchorage, AK (2011)
95. Mei, S., Cai, H., Bu, F.: Multi-view Service-oriented Rule Merged Business Process Modeling Framework. In: Proceedings of the 6th SOSE, pp. 175–180. Irvine, CA (2011)
96. Wang, X.P., Zhou, S.Y., An, Y.F.: Building Flexible SOA-Based Enterprise Process Using Decision Services. In: 7th ICEBE, pp. 270–277. Shanghai (2010)
97. Barjis, J.: Enterprise Process Modelling Complemented with Business Rules. International Journal of Business Process Integration and Management 5, 276 (2011)
98. Hurbean, L.: A New Approach in Business Process Management. München (2008)
99. Schacher, M.: Integrated Enterprise Modelling Based on OMG Specifications. In: Proceedings of the 12th EDOC Conference, xxvii-xxvii. Munich (2008)
100. Object Management Group: Documents Associated with Decision Model and Notation (DMN), Version 1.1, http://www.omg.org/spec/DMN/1.1/ (Abgerufen am: 18.08.2016)
101. Beckmann, J.A.: Business Process Modeling: Software Engineering, Analysis and Applications. Nova Science Publishers Incorporated (2010)
102. Schmiedel, T., Vom Brocke, J., Recker, J.: Culture in Business Process Management: How Cultural Values Determine BPM Success. In: Vom Brocke, J., Rosemann, M. (eds.) Handbook on Business Process Management 2: Strategic Alignment, Governance, People and Culture, pp. 649–663. Springer, Berlin (2015)
103. Theresa Schmiedel, Jan vom Brocke, Jan Recker: Development and Validation of an Instrument to Measure Organizational Cultures' Support of Business Process Management. Information & Management 51, 43–56 (2014)
104. Bjoern Niehaves, Jens Poeppelbuss, Ralf Plattfaut, Joerg Becker: BPM Capability Development: A Matter of Contingencies. Business Process Management Journal 20, 90–106 (2014)
105. Malinova, M., Hribar, B., Mendling, J.: A Framework for Assessing BPM Success. In: Proceedings of the 22nd European Conference on Information Systems (ECIS), pp. 1–15. Tel Aviv (2014)
106. Schmiedel, T., Vom Brocke, J., Recker, J.: Which Cultural Values Matter to Business Process Management? Business Process Management Journal 19, 292–317 (2013)

Detecting Compliance with Business Rules in Ontology-Based Process Modeling

Carl Corea[1], Patrick Delfmann[1]

[1] University of Koblenz-Landau, Institute for IS Research, Koblenz, Germany
{ccorea,delfmann}@uni-koblenz.de

Abstract. Extending business processes with semantic annotations has gained recent attention. This comprises relating process elements to ontology elements in order to create a shared conceptual and terminological understanding. In business process modeling, processes may have to adhere to a multitude of rules. A common way to detect compliance automatedly is studying the artifact of the process model itself. However, if an ontology exists as an additional artifact, it may prove beneficial to exploit this structure for compliance detection, as it provides a rich specification of the business process. We therefore propose an approach that models a rules-layer ontop of an ontology. Said rules-layer is implemented by a *logic program* and can be used to *reason* about the compliance of an underlying *ontology*. Our approach allows ad-hoc access to external ontologies, other than similar approaches that are reliant on a redundant logical representation of process model elements.

Keywords: Compliance Management, Business Rules, Business Process Models, Business Ontologies

1. Introduction

Compliance management is an important part of business process modeling (BPM), aimed to ensure that the company practices which can be entailed from the respective process models are compliant to regulations and business rules [13]. This especially holds for sectors subject to a high degree of regulatory control, such as the financial industry or healthcare [2], [13]. As an example, vendors of financial services may want to warrant that their process for granting loans does not violate any laws or obligations. Compliance management therefore supports *improving* business processes, as potential violations can be eliminated after they have been found, e.g. through re-modeling the business process [3].

The necessity for compliance management has yielded the rise of *automated* approaches, as trying to investigate large company processes for compliance violations *manually* can be seen as an unfeasible task for humans [13]. Following [10], a core notion of such automated approaches is the study of the business process model itself. Such business process models are typically represented through graphical modeling languages such as Event-driven Process Chains (EPC) [26], in order to provide a suitable balance between specification and readability. However, using such languages

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Corea, C.; Delfmann, P. (2017): Detecting Compliance with Business Rules in Ontology-Based Process Modeling, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 226-240

to model business processes, as well as limiting compliance management to this artifact, can bare restrictions in regard to the semantic interpretation of the process model [25]. Although languages such as EPC offer some guidelines towards how to encode process syntax and semantics, the content of the models ultimately lies in the responsebility of the process modeler [21]. Especially when collaboratively creating process models, this can result in different interpretations of the process semantics due to problems such as ambiguity in human language [9], [25]. These different interpretations of the process model can pose potential problems, if they are meant to be analyzed as a central corpus in the scope of compliance management [7], [25].

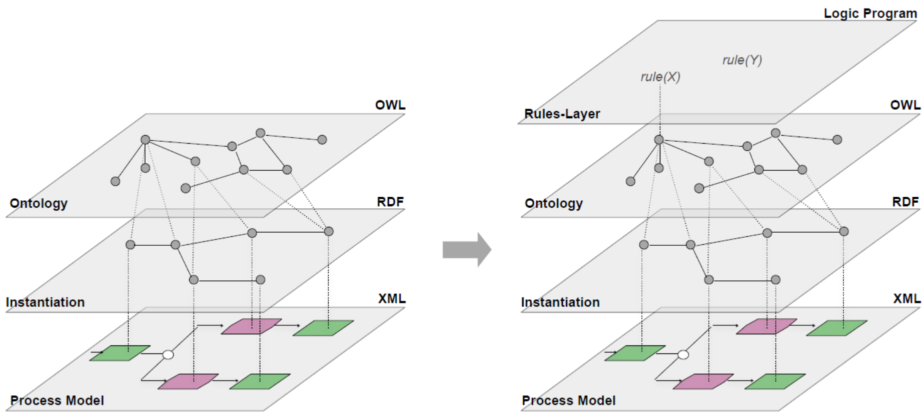


Figure 1. 4-layered framework adapted from [25]

To conquer the problem of different interpretations of business processes, works such as [7-8], [20], [24-25] have proposed to use ontologies to create a shared terminological and conceptual understanding of process models. Ontologies, which Gruber [16] defines as a formal and shared specification of a domain of interest, are a central object of interest in scientific fields such as the Semantic Web [18], which is why works such as [7-8], [20], [24-25] have proposed it may prove as beneficial to investigate applying this object for business process modeling. A main advantage of using ontologies in BPM is that a process model is extended such that machines can access it in a way useful for humans [18]. To this aim, elements of the process model can be annotated to ontology concepts, which is also referred to as *semantic annotations*. This promotes the understanding of the overall business process, as the process model is linked to a conceptual and terminological understanding shared by the modelers. Following [25], ontologies can be furthermore extended by the modelers. For instance, relevant policies or business rules can directly be included in such an ontology, explicitly *specifying the relations between the business process and such regulations*. Concluding, this report investigates utilizing such an ontology as an artifact for compliance management, as an ontology can be considered as an advanced basis for this form of process-oriented compliance management.

This is clarified in figure 1. On the left, which shows the three-layered model for a semantic annotation by [25], we see a business process which has been related to a business ontology. The lower layer represents the classical business process - in this case an EPC diagram. The upper ontology comprises terminological concepts relevant for the respective business as well as their relations. The intermediate instantiation is used to assign elements of the business process to ontology concepts. I.e. instances of the business ontology are used to define the semantics of the EPC model elements. In this paper, we aim at providing an approach which layers a specification of business rules ontop of a business ontology. We refer to our approach as a 4-layered framework, meaning that our contribution extends the existing framework by [25]. This can be seen on the right of figure 1, which shows a rules-layer that can access the underlying ontology. There has been recent attention on combining rules and ontologies in the field of the Semantic Web, due to the complimentary characters of these components [23]. While there is no clear standard yet on how to combine these components, there is a consensus in literature that logic programs can be used to express rules over underlying ontologies [6], [15].

As a result, logic programs can be applied to formalize business rules which can then directly access the vocabulary used in the underlying ontology. Due to the fact that the mentioned ontology is connected to the business process itself, the business rules expressed in the logic program can also access the actual business process itself. Thus, the approach proposed in this report allows to apply the amenities of logic program reasoning in the context of compliance management. By layering rules ontop of already existing business ontologies, detecting compliance is not limited to analyzing the artifact of the business process model itself, but rather the more sophisticated description present in the business ontology can be exploited for verifying compliance in business processes.

The contributions of this report can be summarized as follows. At first, we show that our framework is a novel approach on combining business rules and business ontologies by motivating the 4-layered framework in the context of related work in section 2. Furthermore, after providing a brief recap on how to create semantic annotations for business processes, we show how the business ontology can be integrated into a logic program in order to ensure that company processes adhere to business rules and regulations in section 3. We illustrate our approach and also provide a demonstration in section 4. Finally, our discussion is concluded in section 5.

2. Related Work and Motivation

There have been numerous proposals for automatedly detecting compliance of processes with regulations or business rules [10]. One major school of thought are graph-based approaches [7], [10]. Here, graph-patterns which represent business rule violations are defined o graphically modelled, e.g in approaches like BPMN-Q, eCRG or DMQL[7]. Consequently, a pattern search can be applied to the graph-structure representing the business process in order to find respective violations. It is important to realize that mentioned graph-based approaches focus on analyzing the artifact of the business process model itself.

As mentioned, process models can be linked to business ontologies. In result, next to the artifact of the business process itself, the company may have a second artifact of a business ontology, which can be used as a basis for compliance management. This has been proposed by works such as [7-8][20][24-25], due to the sophisticated semantic structure offered by ontologies. Our approach is therefore an extension to works such as [25], which tries to capture the possibilities that are potentially present.

A main concept of our approach relates to defining business rules as logic expressions. Many others have already proposed using logic expressions instead of graph-based approaches. For example, there is a broad consensus in academia, that temporal logic is suitable to check process models against business rules [10]. While we do not disagree with this claim, we would like to point out that there are limitations of using temporal logic for this aim that have been identified by works such as [10]. For example, investigating process elements with complex annotations or dependencies can be seen as a very complex task [7], [10]. We therefore argue, that using temporal logic for compliance checking should not be taken as self-evident. Authors like Gruhn and Governatori also agree that such formalisms rooted in temporal logic may suffer from some limitations in compliance checking and have therefore proposed other families of logic for this use-case [13], [17].

Said authors have shown that logic programs can be used to validate syntax or compliance in business processes. Following [22], using logic programs to verify business rules is applied as follows: At first, a *new* logic program is derived from a process model. To clarify, the logic program is independent of the original process model. All model elements and their relations are *redundantly* translated into a logic representation. Only then can the logic program reason about compliance. In our approach, we propose to layer a logic program ontop of the already existing process ontology. In this way, process elements do not have to be redundantly translated into a logic representation of an independent logic program, but rather business rules in the form of logic expressions can directly access the underlying ontology and reason about the compliance of the ontology, respectively the business process itself. In case of changes to the business process or the business ontology, our approach is therefore still able to verify compliance without the effort of having to repeat a redundant translation of changed elements.

Works such as [6], [15] have investigated integrating logic programs and ontologies. Said works are however for the Semantic Web and are not specifically aimed at business process management. [8] have also proposed an ontology-driven approach to detect compliance with rules. Here however, they do not investigate using logic programs in order to express regulations and rules.

To the best of our knowledge, our approach is the first to study the intersection between *(a) using rules to detect compliance based on the artifacts of both a process model and a business ontology,* and *(b) using logic programs to implement these rules*. The framework therefore allows to exploit logic program reasoning relative to a business ontology, without having to translate elements of the business ontology into a redundant logical representation in a respective logic program. Table 1 positions our approach in the above mentioned intersection.

Table 1. Approach research gap

Literature	*Uses logic programs to detect compliance*	*Logic programs can access underlying artifacts in an ad-hoc manner*
[1][13-14] [17][22]	x	
[6][12][15] [19][23]		x
Proposed Approach	x	x

3. Layering Rules Ontop of Business Process Ontologies

This section introduces modeling a rules-layer ontop of a business ontology.

3.1. Ontology-Based Process Modeling

Following [25], the scientific results in regard to extending information with semantic annotations can successfully be applied to BPM. Said authors employ a three-layered approach, combining the actual *business process model* with a *business ontology* through an intermediate *instantiation*, as can be seen in figure 1.

The ontology classes define terminological knowledge relevant to the company. For example, *entities* such as organizational units, tasks, events, services or rules and their individual *relations* can be modeled. It is important to realize, that companies do not necessarily have to model such an ontology themselves. Works such as [20] have already proposed reference ontologies that can be re-used and adapted to individual company requirements. In this work, we assume that the ontology is stored in the web-ontology language format (OWL)[1], which is the W3C standard for knowledge representations. Hence, next to the already introduced *terminological* knowledge, *axiomatical* instances of ontology concepts can be created. This is shown in figure 2, which provides an exemplary business ontology. One can observe that this ontology is subdivided into classes (e.g. Unit) and instances (e.g. Production and Sales), whose relationships are defined by this OWL graph-structure.

As a next step, instances of the business ontology are used to define the semantics of process model elements. We want to emphasize that this approach can be applied to arbitrary process models [25].

A process model is defined by element types and specific elements of such types [4]. The ontology can therefore be used to define the semantics of element types, also referred to as *language constructs,* and elements, also referred to as *model elements*. Language constructs, such as events, functions or connectors, should be modeled in the

[1] https://www.w3.org/TR/owl2-syntax/

reference ontology as can be seen in figure 2. As mentioned, the language constructs represented in OWL can be furthermore extended to fit individual company needs. For instance, EPC events could be specialized to create unique and distinguishable event types relevant to the company. Next, model elements can be represented in the ontology through an instantiation of the previously defined language construct classes. Figure 3 illustrates mapping an EPC diagram to a business ontology excerpt. As can be seen, the ontology comprises all relevant EPC language constructs, represented through the respective ontology classes. Every model element resided in the business process model is assigned to an ontology instance. In this way, ontology instances represent the individual model elements from the viewpoint of the business ontology [25]. As a main result, this approach has generated a conceptual viewpoint for the EPC diagram by making the process model accessible in the ontology. This viewpoint could already be exploited to pose conceptual queries regarding the business process. Assuming an OWL ontology, the W3C query language SPARQL[2] can be employed to answer such queries [18].

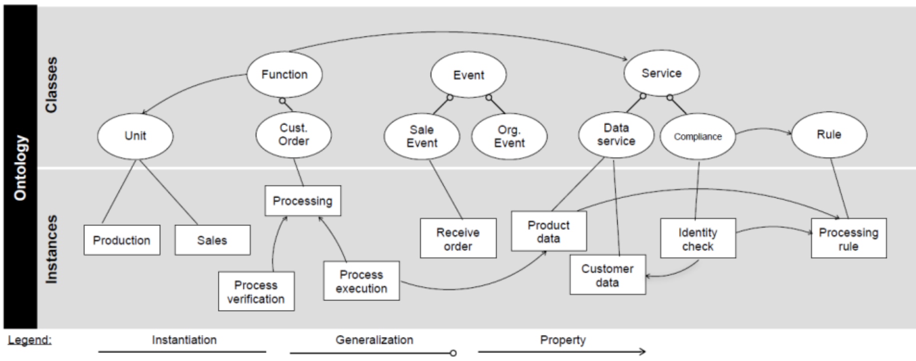


Figure 2. Exemplary business ontology

The model elements of the EPC diagram can now be assigned to further ontology concepts in order to create richer semantic annotations and therefore foster the semantic understanding of the business process, as individual model element semantics and their relations can be explicitly defined. Following [25], a major advantage of this approach is that copious constructs offered by the OWL formalism can be exploited to define formal semantics of great granularity. For example, the possibility to model generalizations or properties such as transitivity, symmetry or inversion between OWL instances allows to use sophisticated reasoning capabilities to analyze process models.

Business ontologies can be modeled or extended according to company needs. Consequently, business rules or regulations and their relation to the business process can be incorporated in the ontology. In result, the business ontology can be seen as an advanced artifact to use in the scope of compliance management [7], [25]. This leads us to our proposal, namely to extend the framework by [25] by modeling a rules-layer ontop of the ontology-layer. This rules-layer should offer the possibility to express

[2] https://www.w3.org/TR/rdf-sparql-query/

business rules and regulations relative to the business ontology and verify their compliance accordingly. To this aim, we propose to utilize a logic program formalism in order to implement said rules-layer, which we introduce subsequently.

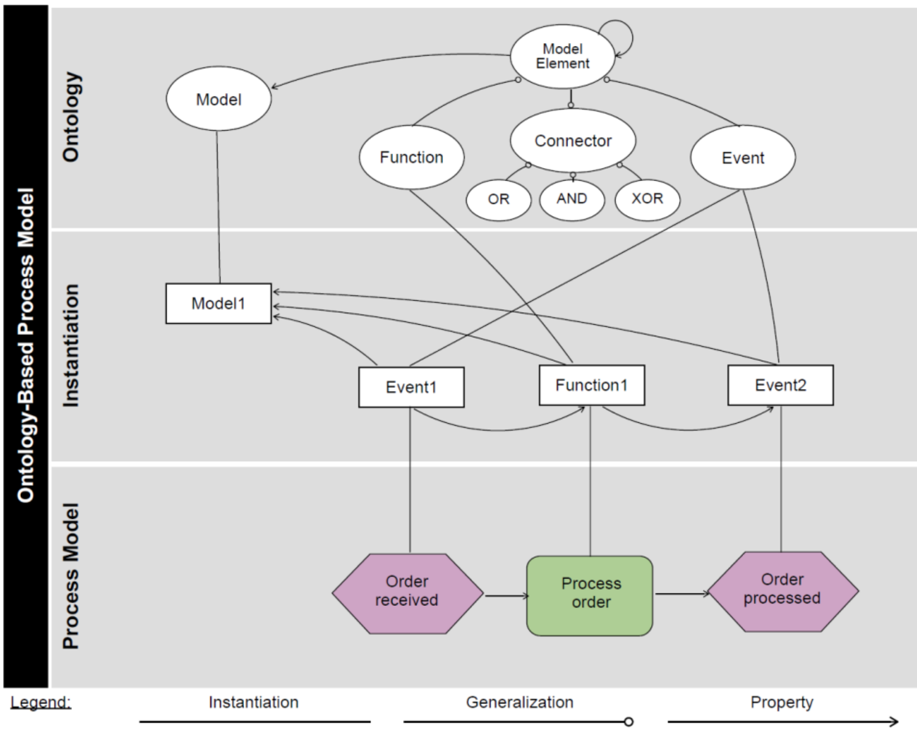


Figure 3. Mapping an EPC diagram to an ontology

3.2. DL-Programs

The Semantic Web architecture allows to model rules ontop of knowledge representations, i.e. ontologies [18]. In this way, the complimentary characters of these two layers can be used to promote automated processing mechanisms. As an important design choice, the rules-layer and the ontology-layer should be interoperable but abstracted from each other, as dictated by the Semantic Web architecture [18]. While there is no clear standard yet on which technologies to use for combining rules and ontologies, there have been several proposals [6].

What is important in the context of this work, is that research suggests that the mentioned rules-layer can be implemented by logic programs [15]. To clarify, logic programs can be used to express rules that mention the vocabulary of an external ontology, i.e. a description logic knowledge base. Thus, the sophisticated reasoning possibilities offered by logic programming can be used to infer information regarding ontologies - more specifically business ontologies. Again, as there are many families of logic, many proposals have been made as to which form of logic programs to use in order to express rules for underlying ontologies. For a detailed survey, please see [15]. As a design choice, we have chosen an approach by [6] entitled *DL-programs*. DL-programs are a formalism able to extend logic programs with description logic expressions. More specifically, they allow to combine normal programs and description logic ontologies, respectively answer-set semantics with first-order semantics. Therefore, they meet our requirement of being able to express business rules and regulations relative to a business ontology.

Recap of Logic Programming. To recall logic programs in general, a logic program is defined as a tuple $\boldsymbol{t} = (\mathbf{P},\mathbf{C})$, where **P** is a set of predicate symbols and **C** is a set of constants [23]. The rudiments of this signature can be used to express *rules*. Such a rule *r* consists of a *premise* and a *conclusion* of the general form

$$\mathtt{head} \longleftarrow \mathtt{rule} \tag{1}$$

meaning that the head, or conclusion, of *r* is true, if the body of the rule is satisfied [23]. If the body of *r* is empty, *r* is referred to as a simple fact. The head and body of the rule can be used to compose formulas of the form

$$h \longleftarrow a_1, \ldots, a_n, \mathit{not}\ b_1, \ldots, b_m \tag{2}$$

where each h, a_i and b_i are so-called *atoms*. Such an atom is defined as $p(t_1, \ldots, t_m)$, where p is a predicate symbol of **P** and every t_i is either a constant from **C** or a simple variable, the latter denoted by a capitalized character. Note that every a_i is a so-called positive atom, and every b_i is a so-called negative atom, indicated by the *not*.

As an example, the following logic program in figure 4 could be used by a financial service to define rules regarding account values.

```
(i)     account(a, 100)
(ii)    account(b, -100)
(iii)   positveBalance(A) ⟵  account(A,B), B > 0
(iv)    error(A) ⟵  account(A), not positiveBalance(A)
```

Figure 4. Exemplary logic program

The first two lines of this logic program are simple facts, stating two accounts entitled *a* and *b* and their respective account balance. The rule in (iii) is used to verify if an account has a positive balance. Here, the capital *A* and *B* represent variables, respectively the account name and account balance. The body of this rule, i.e. the premise, is satisfied if *B* is positive, meaning that only in this case *positiveBalance(A)* could be concluded. A customer account may be required to be of positive balance due to business requirements. Hence, the rule in (iv) models a violation of this requirement. The head of this rule - *error* - is true as soon as there exists and account which is not of *positiveBalance*. As the account named b represents such a case, the depicted logic program can be used to entail that an error is present.

Answer-Set Semantics. The use of variables allows to entail that an error is present for account *b*. This is an example of so-called *answer-set* semantics of logic programs [23]. In early research on logic programs, variables were not included, meaning that rules consisted of simple forms similar to a ⟵b. Such logic programs can be used to entail so-called *well-founded* semantics [11], which can be seen as simple *proofs*. Other than such simple proofs, a paradigm shift to answer-set semantics, which can be traced back to the works of Gelfond and Lifschitz [11], has allowed to define *answer-sets*. These answer-sets can be understood as a model satisfying a logic program. Consequently, an answer-set $M = \{ \ account(b,-100) \ \}$ can be derived from the exemplary logic program. As a result, the financial service could benefit from such an answer-set to identify specific accounts violating the business rule regarding a positive account balance [11].

Introduction to DL-Programs. Continuing our exemplary logic program, computing an answer-set was limited to facts and rules *contained* in this logic program. In order to allow the logic program, i.e. the rules, to reason about information in external ontologies, they have to be extended in such a way that they can *access* said external knowledge bases.

DL-programs [6] represent such an extension of logic programs that allow to access vocabulary of underlying ontologies. A DL-program consists of a normal program *P* and an external knowledge base, i.e. ontology, *L* [6]. While *P* is a finite set of rules based on predicates and constants as introduced, the ontology comprises concepts, roles and individuals. In result, such a DL-program can use a business ontology as a knowledge base *L* and layer a logic program *P* ontop of it.

A question that may arise is how exactly *P* can be extended to access concepts, roles and individuals in *L*. To this aim, [6] have proposed to extend the rules of *P* with so-called *DL-atoms*. While it is not our intent to elaborate on the syntax of such DL-atoms in great detail, it is sufficient to realize that DL-atoms are of the general form:

$$DL[Q](t) \tag{3}$$

Inspecting this complex more closely, the sequence *DL* signalizes the beginning of a DL-query. One can observe that such a DL-query consists of *Q*, which may refer to a *concept* or *role* of the knowledge base *L*. The *(t)* is a simple logic program *term* as introduced earlier, i.e. a constant or a variable. By extending the logic program *P* with the DL-query *DL[Q]*, it can refer to a vocabulary *Q* of an external ontology. As an example, figure 5 shows a DL-program based on the following logic program *P* and knowledge base *L*:

```
L:
(i)         event ⊑ modelElement
(ii)        event(e1)
(iii)       event(e2)

P:
(iv)        evt(X)  ⟵ DL[event](X)
```

Figure 5. Exemplary DL-program

In this example, a business ontology *L* contains terminological knowledge about the language construct *event*, as well as axiomatical *event* instances. The logic program rule in (iv) shows how a DL-atom is used to extend a logic program rule. It is important to realize, that *event* in square brackets of the DL-atoms in (iv) refers to the *event* concept of the business ontology L. I.e. knowledge contained in the ontology L does not have to be *redundantly* expressed as a fact in P, but rather the DL-program allows a logic program P to access an ontology L in an *ad-hoc* manner. To conclude, regarding the logic program P, an answer-set $M = \{ evt(e1), evt(e2) \}$ could be directly entailed by the means of using (iv) to match the variable X in the head of the rule with the variable X in the DL-atom. This answer-set could then be processed in additional rules of P.

As can be seen in the example, DL-programs allow logic programs to access the information stored in an external business ontology. This facilitates powerful and expressive ways to process knowledge bases by the means of rule bases formalized through logic programs [6]. In our opinion, enabling a rules-layer to access a business ontology in an ad-hoc manner is a stronghold of our approach which we discuss consequently.

4. Compliance Checking Approach

This section demonstrates how DL-programs can be used to express business rules in order to detect compliance based on both the artifacts of the business process model as well as a business ontology.

4.1. Framework Architecture

Creating an Ontology-Based Process Model. Figure 6 provides an overview of the framework architecture. The proposed approach extends the three-layered model by [25]. On the left, the process model is connected to a business ontology by instantiation. We denote this as an *ontology-based process model*.

Next, the rules-layer consists of a logic program expressing business rules. The specific business rules may originate from business requirements or external regulatory policies. The rules-layer can directly access the ontology using the DL-program formalism. The rules of the logic program can therefore infer information about the entire ontology-based process model, as the ontology is connected to the process model. The proposed approach is therefore capable to detect whether a process model complies with business rules. This is depicted by the compliance detection component. Here, the DL-program is applied to infer answer-sets of compliance violations. To clarify, these answer-sets consist of *ontology-instances* violating rules of the logic program. These answer-sets therefore also reflect specific *process model elements* violating these logic program rules, respectively business rules. These sets of process model elements can consequently be browsed by the modeler and remodeled according to rules and regulations.

A barrier for the implementation of our approach is the necessity for the artifact of an ontology-based process model. Companies must annotate their processes to an ontology. While a company could perform this task manually, there are existing approa-ches showing that this task can be supported automatically [5]. In [5], the authors show that identifiers of processes and ontologies can be terminologically standardized and thus matched accordingly. This lowers the effort that has to be invested by companies. Undoubtedly, the initial creation of an ontology-based process model has to be considered by companies. However, in our opinion, modeling company processes geared towards a business ontology helps to create a shared understanding across the entire organization. On the long-term, this can be seen as beneficial for the scalability and maintenance of business processes [8].

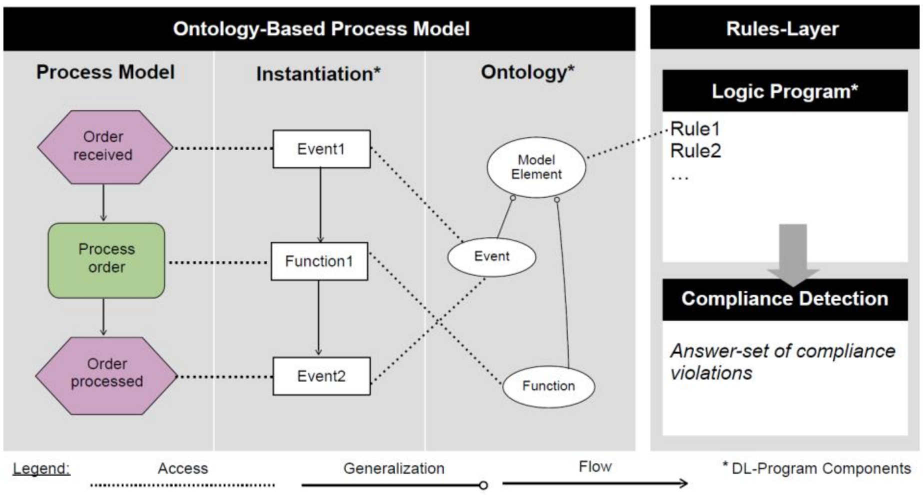


Figure 6. Framework architecture

Expressivenes. While temporal logic can be used to verify compliance, it still suffers from limitations [13], [17]. The expressiveness of our approach therefore aims to conquer some of these limitations, while not sacrificing any amenities. This is achieved through the instantiation. As shown in figure 6, every process model element is represented by an ontology instance. All flows between process model elements are also captured in the instantiation. This means, the execution semantics of the process is encoded in the ontology and can be processed accordingly. In result, sequences, loops or gateways can also be processed by our approach. Van der Aalst et al. [27] have categorized different types of compliance rules. It is beyond the scope of this report to discuss this categorization, but it is used here to specify the expressiveness of our approach in relation to said categorization. So far, we have successfully implemented rules of the categories *existence, precedence, chain precedence, response, between, exclusive, mutual exclusive, inclusive, prerequisite* and *corequisite*. A clear limitation are cardinality restraints or parallel processes. For further details on the DL-formalism, please see [15], as this paper introduces the syntax of DL-queries that are the foundation of compliance checking in our approach.

4.2. Demonstration

To demonstrate our approach, the following exemplary scenario was implemented. We envision a scenario where a company wishes to apply our approach to ensure a business process complies to a business rule. For simplicity, rules and business process will be kept minimal. We assume that the company conducts the task of *paying a bill* within their process. It is furthermore assumed, that a corresponding business rule demands that during this process the bill is *checked* before it is paid. Assuming that the exemplary company aims to create ontology-based process models, figure 7 depicts artifacts which can be utilized in the scope of compliance management for our scenario. Given that a modeler has created the process model in figure 7 (i), our approach allows to detect whether this model complies with the mentioned business rule as follows.

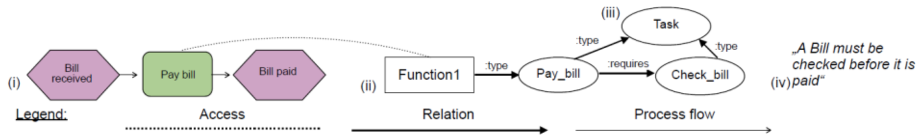


Figure 7. Artifacts available for the compliance management scenario

In (iii), the ontology is depicted as an OWL graph-structure. The company has modeled an ontology class *Task*, which is specialized into *Check_bill* and *Pay_bill*. It is important to realize that these are not instantiations, but simply a specification of the two concepts, which are *tasks*. An instantiation is shown in (ii). The ontology in (iii) was extended by a business rule, indicating that the concept of *Pay_bill* requires *Check_bill*. In this context, the edge labelled *requires* encodes the business rule in (iv). Intuitively, the process model in (i) does not comply to the company policy in (iv), as the bill is not checked before it is paid. The company can implement a DL-program based on a logic program encoding (iv) and the business ontology in (ii)-(iii). Figure 8 depicts this DL-program. For simplicity, namespaces for standard W3C vocabulary such as *OWL:class* are omitted. In the ontology, the class of a *Task* is defined. This is performed analogously for *Check_bill* and *Pay_bill*. Figure 8 also shows how a process model element can be serialized as XML exemplarily. In our scenario, as the modeler has only modeled the task *Pay_bill*, we may only create this single instantiation. The *required* relation, meaning that *Pay_bill* requires *Check_bill*, is also serialized. A logic program can then be layered ontop of this serialization, as shown in figure 8. In the first line of this logic program, an ad-hoc access to instances of the ontology is defined. The *task* in the head of the rule is a logic program predicate, whereas the *Task* in the body of the rule directly refers to the ontology concept by the means of the DL-atom. Line two of the logic program shows a specification of the aforementioned business rule. Here, we can *conclude* an error, if there exists an instance *X* requiring an instance *Y*, and there is no such instance *Y*. The DL-atom in the body of the rule also directly mentions the *required* vocabulary of the ontology. Thus, this relation already present in the ontology can be accessed directly as opposed to being redundantly represented in the logic program. In result, the logic program can correctly infer that an error is present, through the answer-set *M = { error("Pay_bill", "Check_bill") }*. Such inferences can not only be drawn theoretically, but also by many logic program reasoners such as RACER or Hermit [7]. [6] provide a web-interface[3] allowing to enter DL-programs. The serialization as shown in figure 8 can be entered in this web-interface in order to

[3] https://www.mat.unical.it/ianni/swlp/

conclude that mentioned error is present. In our opinion, this shows that our approach can successfully be applied to implement our scenario, namely to detect wether a process model complies to a business rule. Given the artifact of an ontology-based process model and a logic program encoding rules and regulations, the technology to conduct compliance management on the basis of our approach is ready to use.

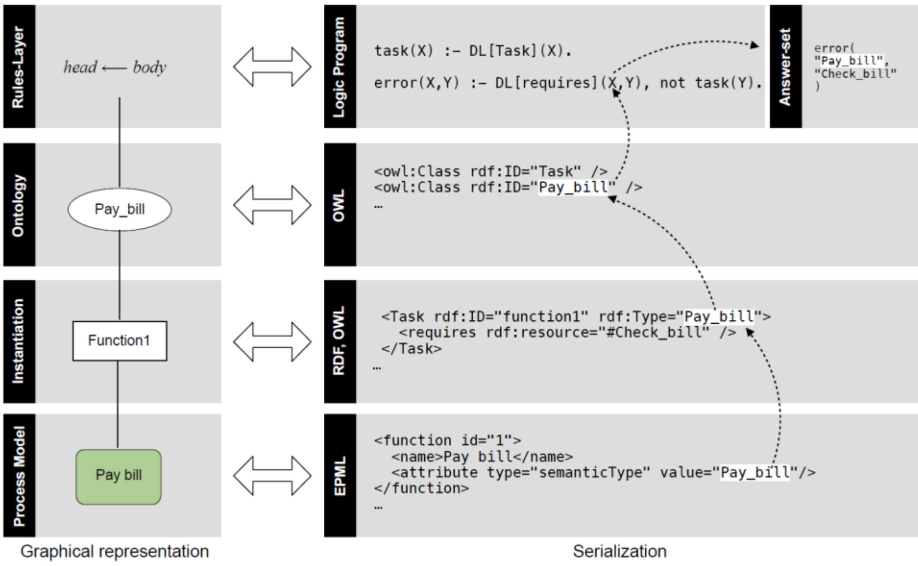


Figure 8. DL-program to entail process model elements that violate business rules

Our example covered a rule that checks existence. This can be extended to model different types of compliance rules, as mentioned in the subsection on expressiveness. For example, one could verify that the *Pay_bill* function is executed between two events, or that a function A is limited in precedence over a function B.

5. Conclusion and Outlook

The contribution of this paper is an approach capable of verifying if a business process complies with business rules, based on analyzing the artifact of a so-called ontology-based process model. The proposed approach allows to (a) specify business rules as logic program expressions relative to an external business ontology, (b) utilize logic program reasoning to find process model elements violating these rules and (c) access information stored in the business ontology directly, i.e. without a redundant transformation of ontology-instances into a logic program representation.

An exemplary implementation of our approach shows that our approach can be successfully applied to find sets of process model elements violating business rules. Future work is to be directed to apply our approach to large-scale process models and business rules. However, the success of using DL-programs for Semantic Web datasets [6] leads us in our belief, that applying our approach to business related data-sets should pose no significant computational problems per se.

A clear limitation of our approach is, that it is dependent of (1) a business ontology connected to a process model and (2) a logic program rule base. Following [7], implementing these artifacts is not yet significantly performed in practice. Literature however strongly suggests the potential of using these artifacts for process modeling and compliance management [7], [13], [20], [22], [24-25]. We therefore see great research potential in assisting companies to create and manage these artifacts. Works such as [5] show that these tasks can be supported automatedly.

As a conclusion, incorporating semantics in the scope of compliance management can contribute towards finding violations in business processes and therefore aid the improvement of company process. Here, using logic program techniques to *reason* about business *ontologies* assists the automated detection of compliance violations. Our approach, allowing an ad-hoc access for rules relative to a process model, lowers the effort that has to be directed towards this aim by companies.

References

1. Antoniou, G., & Arief, M. (2001). Executable Declarative Business rules and their use in Electronic Commerce. Intelligent Systems in Accounting, Finance and Management, 10(4), 211-223.
2. Becker, J., Delfmann, P., Dietrich, H. A., Steinhorst, M., & Eggert, M. (2016). Business process compliance checking–applying and evaluating a generic pattern matching approach for conceptual models in the financial sector. Information Systems Frontiers, 18(2), 359-405.
3. Delfmann, P., & Höhenberger, S. (2015). Supporting Business Process Improvement through Business Process Weakness Pattern Collections. In 12th International Conference on Wirtschaftsinformatik (pp. 4-6).
4. Delfmann, P., Steinhorst, M., Dietrich, H. A., & Becker, J. (2015). The generic model query language GMQL-Conceptual specification, implementation, and runtime evaluation. Information Systems, 47, 129-177.
5. Delfmann, P., Herwig, S., & Lis, L. (2009). Unified enterprise knowledge representation with conceptual models-Capturing corporate language in naming conventions. ICIS 2009 Proceedings, 45.
6. Eiter, T., Ianni, G., Lukasiewicz, T., Schindlauer, R., & Tompits, H. (2008). Combining answer set programming with description logics for the semantic web. Artificial Intelligence, 172(12), 1495-1539.
7. Fellmann, M., Delfmann, P., Koschmider, A., Laue, R., Leopold, H., & Schoknecht, A. (2015). Semantic Technology in Business Process Modeling and Analysis. Part 1: Matching, Modeling Support, Correctness and Compliance. EMISA Forum, 15–31.
8. Fellmann, M., Hogrebe, F., Thomas, O., & Nüttgens, M. (2011). Checking the semantic correctness of process models. Enterprise Modelling and Information Systems Architectures, 6(3), 25-35.
9. Fellmann, M., & Zasada, A. (2014). State-of-the-art of business process compliance approaches. In Proceedings of the 22nd European Conference on Information Systems (pp. 1-12).
10. Frappier, M., Fraikin, B., Chossart, R., Chane-Yack-Fa, R., & Ouenzar, M. (2010). Comparison of model checking tools for information systems. In International Conference on Formal Engineering Methods (pp. 581-596).

11. Gelfond, M., & Lifschitz, V. (1988). The stable model semantics for logic programming. In ICLP/SLP (pp. 1070-1080).
12. Gómez, S. A., Chesnevar, C. I., & Simari, G. R. (2009). Integration of rules and ontologies with defeasible logic programming. In XV Congreso Argentino de Ciencias de la Computación.
13. Governatori, G., Hoffmann, J., Sadiq, S., & Weber, I. (2008). Detecting regulatory compliance for business process models through semantic annotations. In International Conference on Business Process Management (pp. 5-17).
14. Governatori, G., & Rotolo, A. (2010). Norm compliance in business process modeling. In International Workshop on Rules and Rule Markup Languages for the Semantic Web (pp. 194-209).
15. Grosof, B. N., Horrocks, I., Volz, R., & Decker, S. (2003). Description logic programs: combining logic programs with description logic. In Proceedings of the 12th international conference on World Wide Web (pp. 48-57).
16. Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing?. International journal of human-computer studies, 43(5), 907-928.
17. Gruhn, V., & Laue, R. (2007). Checking Properties of Business Process Models with Logic Programming. In MSVVEIS (pp. 84-93).
18. Hitzler, P., Krötzsch, M., Rudolph, S., & Sure, Y. (2007). Semantic Web: Grundlagen. Springer-Verlag.
19. Kontopoulos, E., Bassiliades, N., Governatori, G., & Antoniou, G. (2011). A modal defeasible reasoner of deontic logic for the semantic web. International Journal on Semantic Web and Information Systems (IJSWIS), 7(1), 18-43.
20. Koschmider, A., & Oberweis, A. (2005). Ontology Based Business Process Description. In EMOI-INTEROP.
21. Nielen, A., Költer, D., Mütze-Niewöhner, S., Karla, J., & Schlick, C. M. (2011). An empirical analysis of human performance and error in process model development. In International Conference on Conceptual Modeling (pp. 514-523).
22. Sadiq, S., Governatori, G., & Namiri, K. (2007). Modeling control objectives for business process compliance. In International conference on business process management (pp. 149-164).
23. Shen, Y. D., & Wang, K. (2011). Extending logic programs with description logic expressions for the semantic web. In International Semantic Web Conference (pp. 633-648).
24. Smolnik, S., Teuteberg, F., & Thomas, O. (2012). Semantic Technologies for Business and Information Systems Engineering: Concepts and Applications. Business Science Reference.
25. Thomas, O., & Fellmann, M. (2006). Semantische Integration von Ontologien und Ereignisgesteuerten Prozessketten. In EPK (Vol. 5, pp. 7-23).
26. Van der Aalst, W. M. (1999). Formalization and verification of event-driven process chains. Information and Software technology, 41(10), 639-650.
27. Van der Aalst, W. M., Ramezani, E., & Fahland, D. (2012). Where did i misbehave? diagnostic information in compliance checking. In International conference on business process management (pp. 262-278). Springer Berlin Heidelberg.

Evaluation of a Pattern-Based Approach for Business Process Improvement

Thomas Falk

University of Regensburg, Department of Management Information Systems, Regensburg, Germany
thomas.falk@wiwi.uni-regensburg.de

Abstract. Although many approaches for business process improvement (BPI) are available and have proven their usefulness, it is often stated that they do not provide sufficient methodological support for all phases of a BPI initiative. To close this gap, a pattern-based approach has been suggested as a suitable means to directly support the "act of improvement" and to provide systematic guidance. However, what is missing is an evaluation that provides well-founded insights into the qualities and benefits of this approach. Therefore, this paper presents the results of a laboratory experiment which was conducted to assess the impact of the BPI-Pattern Approach on a process improvement project. The findings, which are based on a number of hypothesis tests, confirm a positive influence on both effectiveness and efficiency when the BPI-Pattern Approach is used.

Keywords: Business Process Improvement, BPI Pattern, Evaluation, Laboratory Experiment

1 Introduction

Business process improvement (BPI) together with its related methods and techniques remains an important topic in research as well as for practitioners. Among the major drivers prompting companies and other organizations to change their processes are cost savings and increased productivity as well as the need to improve products, customer satisfaction or organizational responsiveness to stay competitive [1].

As a result, there is a plethora of different approaches supporting the improvement or redesign of business processes, e.g., *BPI* according to Harrington [2], *Business Process Reengineering (BPR)* [3], or *Six Sigma* [4], to name but a few well-known examples. Existing approaches provide extensive methodological support for e.g., mapping processes, identifying problems and their root causes, etc. However, when it comes to the actual improvement, i.e., the transformation of a process from its "as-is" to a desired "to-be" state, they often rely on creativity, personal skills and experience for the development of an enhanced process design [5], [6]. This pivotal phase of BPI, the so-called "act of improvement", lacks systematic methodological support [7] as tangible instructions on how to achieve substantial improvements are missing [8]. To address this problem, a BPI-Pattern Approach was suggested that systematically

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Falk, T. (2017): Evaluation of a Pattern-Based Approach for Business Process Improvement,
in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung
Wirtschaftsinformatik (WI 2017), St. Gallen, S. 241-255

supports the “act of improvement” in BPI initiatives [9], [10]. Its core components are a catalog of BPI-Patterns, which convey proven knowledge and instructions of how to achieve process improvements, together with a selection process to guide the identification and selection of appropriate patterns for a given situation. This approach was developed following the design science paradigm [11], [12] and its applicability was already demonstrated by a case study [10]. However, what is missing is an in-depth evaluation of the usefulness and advantageousness of the new pattern-based approach in comparison to conventional practices. Therefore, the paper at hand describes a laboratory experiment, which was conducted to quantitatively evaluate the effectiveness as well as the efficiency of the BPI-Pattern Approach. The results gained from the experiment allow for a better assessment of the potential benefits of using BPI-Patterns for process improvement in practical settings.

The remainder of this paper is organized as follows. Section 2 covers conceptual basics of BPI, the BPI-Pattern Approach, and the evaluation phase in design science research. Section 3 introduces the research model including our hypotheses and describes the design, materials and implementation of the experiment. The results are presented in section 4 and their implications discussed in section 5. Section 6 provides a short summary of the findings and an outlook on further research.

2 Conceptual Basics

2.1 Business Process Improvement (BPI)

The improvement of business processes has become a core task in many organizations and is an integral part of the business process lifecycle [13]. Its overall aim is to make processes more effective, efficient and flexible [2], e.g., by reducing costs or cycle times, and thus to yield competitive advantage. Process improvement is considered as essential for creating sustainable value for customers or innovating products and services [14], and organizations link the improvement of their processes to their business strategies [15]. Existing approaches for process improvement are manifold (see e.g. [2-4], [16]) and have proved their usefulness.

However, it appears that none of them adequately supports the user through all the stages in an improvement project, especially when it comes to the “act of improvement” [7]. This very phase is at the heart of any BPI project where the transformation of a process from its “as-is” state to a desired “to-be” state takes place. Actual instructions on how to improve a process in respect of given objectives (e.g. reduce costs, shorten cycle time, etc.) are very scarce and generic [8], and thus many improvement approaches rely on human creativity and personal experience [17]. Even though one “brilliant idea” can bring about significant improvement, its occurrence is always uncertain and often depends on the know-how of individual key players. The development of substantial improvements remains a major challenge in BPI projects and lacks sufficient methodological guidance [7], [18]. To better support BPI performers, structured methods and techniques are required [19] as well as guidelines describing the necessary changes that have to take place [20]. The use of a pattern-based approach for BPI is one possible solution that meets those requirements.

2.2 BPI-Pattern Approach

The concept of patterns is widely used in IS (e.g., design patterns in software development [21], workflow patterns [22], etc.). In the field of BPI there are several works that address the topic of how patterns can be utilized for the improvement of business processes. One group focuses on the collection and description of generic improvement measures (cf. [23-25]). A second group deals with the question of how to run or organize improvement projects (cf. [26], [27]). Because of their inherent characteristics, the use of patterns is also a promising approach for supporting the "act of improvement". Patterns are suited to precisely describe a working solution for a problem in a specific context [21], which is based on proven knowledge [28], but, at the same time, they are generic enough to be reused in many similar cases [29]. Thus, so-called BPI-Patterns guide through the application of an improvement action and help to overcome the shortcomings of existing BPI approaches as mentioned above.

A BPI-Pattern Approach supporting the "act of improvement" is suggested by Falk et al. [9], [10]. It consists of two main components that are both integrated in a software tool: a *pattern catalog* as a repository of improvement propositions and a step-by-step *selection process* for identifying and selecting suitable patterns. The patterns themselves are derived from BPI literature, case studies or experience from real-life projects [30]. In addition, their applicability and effects have been substantiated by means of a simulation [31]. A standardized template comprising the attributes as defined by the underlying data model [9] ensures a consistent description. Its core attributes reflect the definition of a BPI-Pattern as a reusable *solution* for a certain *problem* in a business process within a certain *context* [9] and comprise instructions on how to change the process' elements (e.g., activities, resources, control-flow, etc.) to transform it from its non-satisfactory "as-is" to a desired "to-be" state. The application of a BPI-Pattern has an *effect* (positive, neutral, negative) on the performance dimensions (cost, time, quality, flexibility). An example of a BPI-Pattern that has a positive effect on cycle time is "Parallelize Activities" where formerly sequentially ordered activities are performed simultaneously. The software tool not only allows for the management of the pattern catalog but also guides the user through the selection process. There are two alternative starting points depending on whether the BPI project is targeted at predefined objectives that should be achieved or driven by problems that are detected in the processes. By matching the effects of the patterns with the goals of the improvement project, only those patterns that show the desired outcome are regarded further. Based on the identification of problems together with a root cause analysis, patterns solving or at least mitigating those particular problems are filtered. Either way, both possibilities require the execution of both steps, i.e. if patterns are first to be selected according to their effects, the next step is to look at the problems and vice versa. Afterwards, the contexts of the patterns are checked against the environmental conditions in the case at hand to make sure that they are applicable in the given situation. Finally, the remaining patterns that are candidates for implementation are prioritized, and a decision has to be made by the user which pattern(s) will be applied in the end. That way, the BPI-Pattern Approach provides a systematic, tool-assisted means to support the "act of improvement".

2.3 Evaluation in Design Science Research

Design science research (DSR) is characterized by the creation of artefacts for a specific purpose that solve practice-oriented problems [11], [32]. The evaluation of those artefacts is an integral part of the DSR methodology [12] and considered vitally important to prove the usefulness of the developed artefacts [33]. To perform an evaluation, a number of strategies making use of various evaluation methods (e.g. interviews, simulation, experiments) have been suggested [34-36]. Experiments as a method are particularly suited for a systematic assessment of an artefact's qualities [37]. For the evaluation of the BPI-Pattern Approach, a laboratory experiment was preferred to other evaluation methods because of the following reasons: In a laboratory environment, confounding variables can be better controlled or be eliminated [38]. In contrast to case studies, e.g. the participants are more homogenous and can be assigned to the groups randomly, the "as-is" process to be enhanced as well as the provided information are exactly the same, etc. In addition, a laboratory experiment shows a relatively high internal validity, which allows to determine if changes of the dependent variables (e.g. improvement effectiveness) result from changes of the independent variables (e.g. use of pattern-based approach) or not [39].

3 Evaluation of the BPI-Pattern Approach by an Experiment

For the evaluation of the BPI-Pattern Approach (see section 2.2) and to answer the question regarding its utility a laboratory experiment is conducted (cf. [37-39]). The focus of the study is on both the effectiveness and efficiency when seeking to improve existing "as-is" business processes. For that purpose, the experiment builds on a comparison of two groups: one group uses the BPI-Pattern Approach to identify potentials for improvement and generate process changes in a guided and structured way whereas the second group (control group) works without a systematic approach but is free in using creativity and innovative ideas to enhance the process (one factor with two treatments). As a general principle a completely randomized design was chosen where the subjects are randomly assigned to the groups. The setup of the laboratory experiment is explained in detail in the following section.

3.1 Experimental Design

Our research model (see Figure 1) is based on the underlying assumption that the use of a systematic improvement approach, such as BPI-Patterns, leads to better results (e.g., cost or cycle time reductions) [20], [25], and reduces the overall effort of identifying such improvement possibilities and developing the new "to-be" process design (i.e. time savings in process improvement projects) [23]. The quality of the results increases when tried and tested patterns are used, a concept which is already successfully used in other fields (e.g. software development) [21]. It seems reasonable that this is equally valid in the context of BPI and experiences so far are promising [23], [24]. Based on these assumptions, the following pairs of hypotheses (see Table 1) can be stated which are to be tested by means of the laboratory experiment.

Table 1. Hypotheses on the effects of the BPI-Pattern Approach

Improvement Effectiveness:	
H_0a:	The use of the BPI-Pattern Approach does not generate better improvement solutions than relying on creativity skills.
H_1a:	The use of the BPI-Pattern Approach does generate better improvement solutions than relying on creativity skills.
Expenditure of time:	
H_0b:	The use of the BPI-Pattern Approach does not reduce the time needed to identify improvement potential and to develop an enhanced "to-be" process design.
H_1b:	The use of the BPI-Pattern Approach does reduce the time needed to identify improvement potential and to develop an enhanced "to-be" process design.
Improvement Efficiency:	
H_0c:	The use of the BPI-Pattern Approach does not increase the time efficiency when developing an improved process design.
H_1c:	The use of the BPI-Pattern Approach does increase the time efficiency when developing an improved process design.

All elements of the research model have to be operationalized for the purpose of the actual experiment at hand. It is based on a single binary, independent variable, which is whether a participant is supported by the BPI-Pattern Approach or not. The dependent variables are 1) *Improvement Effectiveness*, 2) *Expenditure of Time*, and 3) *Improvement Efficiency*. *Improvement Effectiveness* is defined by the extent to that the "as-is" process has been improved; i.e. the process performance of the newly developed "to-be" process exceeds the level of the "as-is" process. For complexity reasons this variable has been narrowed down to the dimension time and considers reductions in the average total processing time per process run (in percent). The instrument used for measuring are the adapted "to-be" process models, handed in by the participants. Total processing time is calculated by adding up the processing times of all process activities applying the probability of execution of the respective process path for weighting. *Expenditure of Time* quantifies how much time (in minutes) is needed to complete the "act of improvement", i.e. the very phase where improvement potential is identified and an enhanced process design is developed. The participants are guided by a road map and have to fill in the precise time when they start or finish a specific task. Finally, the variable *Improvement Efficiency* expresses how much improvement could be achieved per time unit and is calculated as achieved process improvements relative to the time needed for their development.

In addition, the research model covers the following factors which have a possible influence on the outcome: *Methodological Knowledge*, *Domain Knowledge*, *BPI Experience*, *Process Understanding*, and *Problem Detection*. To largely eliminate the influence of those factors, we deliberately chose students taking a particular course in process modelling as subjects. Hence, they have a similar background regarding their theoretical and practical skills in process modelling, analysis and improvement which helps to keep the confounding factors stable. Since the focus of the experiment is on evaluating the effect of the BPI-Pattern Approach alone, it is hereby prevented that the participants' diverging personal experience gain too much influence on the experiment results. The aforementioned factors are measured using a set of questions within a questionnaire. The variables *Methodological Knowledge*, *Domain Knowledge*, and *BPI Experience* are covered by a number of questions (e.g., „How often do you deal with BPI?"), each of them having a four-item, ordinal scale with point values assigned. To gather information on *Process Understanding*, open questions as well as true/false questions about different facts in the "as-is" process are used. To determine the degree of *Problem Detection*, the participants had to fill in a free text field to answer the question "What problems could you identify in the 'as-is' process that may cause long processing times?". The value for this variable is calculated as the portion of problems identified by the participant in relation to all problems that are contained in the process model. Because those variables addressed by the questionnaire have a potential impact on the results of the experiment, it is important that they are controlled or at least made explicit so that they can be taken into account when analyzing and interpreting the results [38], [39].

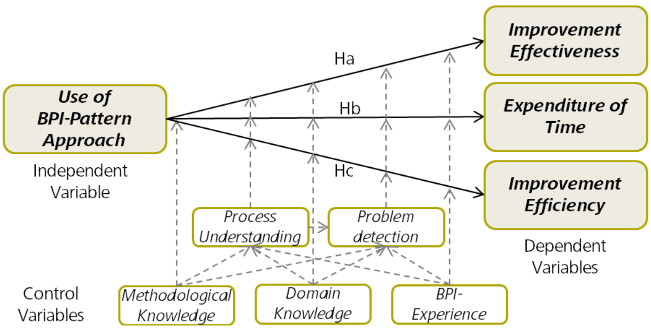


Figure 1. Research model

3.2 Implementation

The experiment was conducted in form of a fictional BPI project working on the matriculation process of a German university that has been slightly adapted for the purpose of the experiment. It has to fulfil certain criteria such as being representative of a practical and realistic BPI case, being sufficiently complex to be resistant to overly obvious or trivial enhancements, but at the same time remaining manageable within the scope of a laboratory experiment. The predefined goal of the project was to alter the given matriculation process such that its processing time is reduced as far as possible but the process still fulfils its core tasks as requested. The participants in the experiment, undergraduate students attending a bachelor's degree course in business process modelling, were randomly assigned to one of two groups: group A used the BPI-Pattern Approach to support the "act of improvement" whereas group B worked without BPI-Patterns. Because the same process model was used for all tests due to comparability reasons, the assignment to the groups was strictly alternative. Otherwise, learning effects while working on the process with one method would have corrupted the results when the second method was used afterwards. In total, 47 students participated in the experiment, 23 in group A and 24 in group B, thus leading to a balanced design with equal group size.

The participants in Group A, who worked with the BPI-Pattern Approach, used a prototypical tool that supports the management of the pattern catalog and especially the selection process of applicable BPI patterns based on certain criteria, such as desired effects, problems to be solved, or contextual requirements (cf. section 2.2). For the purpose and scope of the laboratory experiment the tool was populated with a catalog of 12 different BPI-Patterns. Beyond that, all participants were provided with the same materials: information about the project goals and the measurement of the performance indicator processing time, a textual description of the “as-is” process together with some fixed requirements that cannot be changed, and transcribed excerpts of interviews with the process owner and employees. Furthermore, a process model in eEPC notation was provided both as a paper-based and as a processable electronic version in ARIS Business Architect. The chosen process consisted of 41 process steps and contained several decision points, branches and loops (24 AND/XOR connectors) to express an adequate level of complexity. The process model, besides the graphical representation of the process flow, contained in particular information about the processing time (in minutes) of each activity and the probability values attached to decisions (XOR connectors) during the process operation. In addition, details about the organizational units involved, the requested or created documents, and the required data were provided. The participants’ task was to improve the provided “as-is” process model using the ARIS business architect, by either using the BPI-Pattern Approach or relying on brainstorming and individual creativity skills. There were four major improvement possibilities (e.g. eliminating redundant tasks, re-arranging existing activities in a more logical order, etc.) being supposed to be identified by the participants. Improvement suggestions beyond that, regardless of whether made by group A or B, are also considered as long as they meet the requirements specified in the materials of the experiment. The created “to-be” process models together with the filled in questionnaires have been submitted to the researchers at the end of the experiment.

Guidelines on how to proceed in course of the experiment were provided on a form sheet which instructs the participants and lead them through the experiment step by step. In addition, the procedure was explained by the researchers to the participants before the actual experiment started. Thus, it was ensured that every participant clearly understood the course of action and followed the same standardized steps while working on the fictional process improvement project.

All materials used in the experiment had been pretested before by three fellow researchers as well as a test group of eight students whose state of knowledge compared to that of the actual participants. Both tests provided valuable feedback which was used to refine the concerned materials and/or instruments and confirmed that the provided information, questions and tasks are clear and understandable.

4 Results

Subsequent to the experiment, the 47 submitted “to-be” process models and the questionnaires were carefully examined by two researchers independently. In so doing, each process model was double-checked by a fellow researcher; first without knowing the results from one another to avoid any bias, afterwards, the results were compared against each other to settle different interpretations. Four of them had to be excluded because they were either incomplete, showed data inconsistencies, or allowed for diverging interpretations. Hence, 43 applicable results, 21 in group A and 22 in group B, remain for further analysis and are described in the following.

Table 2. Results of the experiment: values of the dependent variables

Variable	Group	N	Mean	Std. Deviation	Std. Mean Error
Improvement Effectiveness	A: with BPI-Patterns	21	.214467	.1114372	.0243176
	B: without BPI-Patterns	22	.142259	.1136285	.0242257
Expenditure of Time	A: with BPI-Patterns	21	27.05	8.680	1.894
	B: without BPI-Patterns	22	28.27	6.734	1.436
Improvement Efficiency	A: with BPI-Patterns	21	.00810481	.000913604	.000913604
	B: without BPI-Patterns	22	.00512436	.000940070	.000940070

Since we are interested in the effect of the BPI-Pattern Approach in isolation, it is an important question if the observed effects on the dependent variables result from changes of the independent variable or are otherwise attributable to external, confounding factors. Therefore, as a first step, we analyzed the values of the control variables for both groups. Altogether, the two groups in the experiment were quite homogeneous; both, the mean values and the distributions were almost identical. This is also the reason why no statistically significant correlation between any of the control variables and the dependent variables could be verified. A regression analysis based on models that contain the control variables only, shows no significant results (F-test with p-values between .078 and .673) and reveals that such models cannot provide a reliable explanation for the changes observed in the values of the independent variables. Including “Use of BPI-Pattern Approach” as an independent variable increased the model quality as explained further below. The model assumptions, e.g. Gauss-Markov theorem, are sufficiently satisfied; the residues are independent and approximately normally distributed. Based on the aforementioned points, we can fairly assume that the changes we observed, e.g., for effectiveness or efficiency, are referable to the BPI-Pattern Approach. Both foster a high internal validity of the experiment’s results.

To evaluate the impact of the BPI-Pattern Approach on business improvement initiatives, the values for the variables of interest, namely Improvement Effectiveness, Expenditure of Time, and Improvement Efficiency, were measured. Hereafter, the mean values of both groups are compared to determine the effects of the BPI-Pattern Approach. Table 2 shows the values of the arithmetic mean, the standard deviation, and the standard error of the mean for groups A and B, respectively. To test our three hypotheses, which we posed at the beginning of section 3.1, a two-sample t-test was used. According to a Kolmogorov-Smirnov test, the variables are normally distributed for each group. The results of these hypothesis tests are shown in Table 3. In addition, a nonparametric Wilcoxon-Mann-Whitney test, which is more robust than the t-test, was conducted and shows the same results.

Table 3. Hypothesis test for Ha, Hb, and Hc

Variable	t-test for Equality of Means					95% Confidence Interval of Diff.	
	t	df	Sig.	Mean Diff.	Std. Err. Diff.	Lower	Upper
Improvement Effectiveness	2.103	41	.042	.0722076	.0343413	.0028540	.1415612
Expenditure of Time	-.518	41	.607	-1.225	2.363	-5.997	3.547
Improvement Efficiency	2.271	41	.028	.002980446	.001312497	.000329805	.005631087

Improvement Effectiveness is represented by the reductions in processing time comparing the new "to-be" process design with the original "as-is" process. The mean values for the degree of improvement are approx. 0.21 and 0.14 for group A and B, respectively, whereas the standard deviation shows similar values with group A lying slightly below group B. The results show a significant increase of the improvements that were achieved when the BPI-Pattern Approach was used during the experiment. The significance level is above 95% and the lower and upper bound of the 95% confidence interval show the same sign. This leads us to reject the null hypothesis H_0a and to support our assumption that the use of the BPI-Pattern Approach increases the improvement effectiveness, i.e. it generates better improvement results than by relying solely on e.g. creativity skills. The regression model including the independent variable as well as control and moderator variables is significant at p=.039, with "Use of BPI-Pattern Approach" having the strongest influence (p=.026 on "Improvement Effectiveness".

Regarding the second hypothesis, the use of the BPI-Pattern Approach would reduce the time that is needed to develop an enhanced "to-be" process design, the results are inconclusive. Based on the p-value of .607, the null hypothesis cannot be rejected. Moreover, the lower and upper bound of the confidence interval change signs. Therefore, no clear statement can be given about the influence of the BPI-Pattern Approach on the expenditure of time that is needed to improve a business process. However, the average time span needed for the "act of improvement" was slightly lower in group A than in group B (27.05 vs. 28.27 min) and the bigger part of the 95% confidence interval lies in a negative range. Moreover, the regression analysis regarding the variable "expenditure of time" resulted in a negative coefficient (β = -1.537) for "Use of BPI-Pattern Approach", indicating, even if not statistically significant, that possible time savings due to the use of patterns can still be assumed but are subject to further scrutiny.

Improvement Efficiency is addressed by the third hypothesis and expressed by the improvements achieved in relation to the time needed for their development. The observed results reveal a significant relationship between the use of the BPI-Pattern Approach and the variable improvement efficiency, which is on average 1.6 times higher for group A using the patterns ($\approx$ 0.008) as compared to group B not using the patterns ($\approx$ 0.005). Both the signs of the lower and the upper bound of the confidence interval are positive and, thus, confirm the direction we expected. Again based on a two-sample t-test, which reaches a significance level higher than 97%, we reject the null hypothesis H_0c and conclude that the use of the BPI-Pattern Approach increases the time efficiency when developing an improved process design, i.e. “more” improvement can be achieved per time unit. Moreover, the improvement efficiency is more reliable when patterns are used since the standard deviation in group A is lower than those in group B. The regression model for improvement efficiency is even highly significant (p=.004) whereby the two variables Use of BPI-Pattern Approach (p=.009) and Methodological Knowledge (p=.012) show a significant influence.

To get a deeper understanding of the results as well as to point out some implications related to the practical usefulness, they are discussed in more detail in the following section.

5 Discussion

The purpose of our laboratory experiment was to thoroughly evaluate a BPI-Pattern Approach in accordance with the DSR methodology and to determine its capabilities in supporting the “act of improvement” in a process improvement project. The overall results support the conclusion that the pattern-based approach possesses some advantages as all three dependent variables covered by the experiment tend towards the better values in case the BPI patterns are used (see Figure 2). The positive influence on two of them, namely improvement effectiveness and improvement efficiency, is statistically significant in addition. However, the effect size becomes more relevant at a practical level and can only be judged considering the particular characteristics of the business process which has to be improved. For example, a 1%-reduction in time or cost for a mass production process may result in a major impact whereas for other processes with just a few instances the changes have to be many times greater to be of practical relevance. In our case, the average difference in improvement effectiveness between the two groups equals approx. 61 man-days and, thus, has indeed a material impact. In addition, the business process that has to be improved may be an important factor which influences the performance of the BPI-Approach, e.g., based on its structure, complexity, or the type of problems that have to be solved.

The use of the BPI-Pattern Approach increases the improvement effectiveness, which was confirmed by the statistically significant results of the experiment. Taking a closer look at effectiveness, the use of patterns is no guarantee that significant improvements or any improvement at all can be achieved. In our experiment, we found cases where the processing time could not or barely be reduced in both groups even if the participants had applicable patterns at their disposition (see Figure 2(1)). In addition, the case showing the highest degree of improvement occurs in group B, without patterns. A possible explanation is that the BPI patterns primarily aim for incremental, evolutionary enhancements whereas the control group was free to develop more radical improvements, too. This effect could even strengthen when, unlike in the experimental setting used here, teams are assigned to develop the “to-be” process design. In this case, the group dynamics when creating improvement solutions will have some influence on creativity and thus on the improvement effectiveness.

However, what could be shown for the group which used the pattern-based approach is, that if a problem in the "as-is" design of the process was detected correctly, the probability that this problem was solved or at least partly addressed was much higher than in the control group working without patterns. This indicates that the BPI-Pattern Approach significantly increases the chances of finding and successfully implementing a solution, presumed such a one exists.

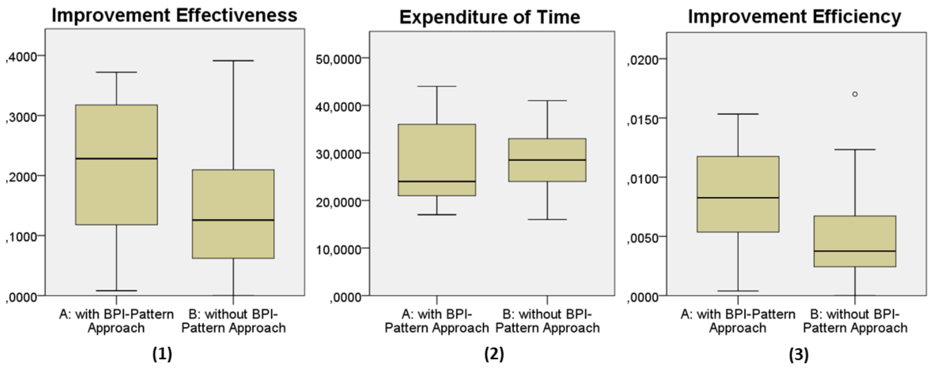


Figure 2. Box plot diagrams of the three dependent variables for group A and B

Even though the arithmetic mean (A: 27.05 min; B: 28.27 min) and the median (A: 24 min; B: 28.5 min) of the variable Expenditure of Time are slightly lower when the pattern-based approach is used, the effect is far weaker than anticipated and thus not statistically significant (see Figure 2(2)). In addition, not a single participant from group A (with patterns) passed through the "act of improvement" in less time than the fastest of group B (without patterns). Similarly, the maximum time in group A exceeds those in the control group. This indeed contradicts our initial expectations, but there are several explanations for these findings. On the one hand, the BPI-Pattern Approach was quite a new method the students were not trained on (participants had obtained a short introduction the week before the experiment started). Because of learning curve effects the results may change when the approach is used over and over. On the other hand, it is easily conceivable that participants who had no BPI-Patterns provided "gave up" if they could not find an adequate solution within a certain time interval. In contrast, the other group continued searching the pattern catalog for patterns that fit their needs.

It is often argued that structured guidelines and a formalized procedure, both provided by the BPI-Pattern Approach, would facilitate process improvement and make redesign projects more efficient [23], [24], [40]. This effect was also clearly detectable in the results of our experiment (see Figure 2(3)). The regression model explaining the dependent variable improvement efficiency shows a high significance which can be ascribed to the BPI-Pattern Approach and the degree of methodological knowledge of the user. What is different to the theoretical considerations in literature is that this is not mainly due to time savings but rather because of implementing more effective improvements. Another topic, discussed by researchers, is that a systematic improvement approach almost necessarily leads to some kind of improvement if only sufficient time and effort is invested [41]. This can be approved based on our observations as we found a significant positive correlation between time and improvement within group A (with BPI-Patterns), but not for the control group. The findings indicate that if one searches long enough some kind of improvement solution will be found, presumed the pattern catalog contains suitable patterns which fit the problem.

Business process improvement projects are often dominated by external advisors who play the role of method experts and contribute the methodological knowledge that is inevitable to successfully run such a project. However, it is recommended to better involve the concerned employees to exploit their domain-specific knowledge as well as to minimize the resistance to change which is a common phenomenon in many organizations [27]. Pattern-based approaches are seen as an instrument to especially support novices by providing step-by-step guidance and conveying expert knowledge they can build on. The data gathered during the experiment clearly classify the participating bachelor students as BPI novices showing relatively low values for Methodological Knowledge, Domain Knowledge, and BPI Experience. However, the overall results of our evaluation show that even such novices in BPI can achieve considerable improvements based on the BPI-Pattern Approach.

The findings of our research revealed insights into strengths and weaknesses of the analyzed approach, too. Examining the data gathered during the experiment we identified possibilities to further improve the BPI-Pattern Approach. First, a number of participants in group A correctly identified the relevant process performance problems but could not find the corresponding match in the pattern catalog. Thus, the problem selection process and the problem description template should be better aligned. Second, those patterns providing more than one example to explain their functional principle were rated to be easily comprehensible. Adding a reasonable number of examples (e.g. covering different application domains) would probably advance the identification and adoption of the patterns.

6 Conclusion

In this paper we present the findings of a laboratory experiment which was conducted to evaluate a BPI-Pattern Approach according to the principles of design science research. As a result, it could not only be demonstrated that the approach works in the particular setting provided in the experiment, i.e. the design science artefact fulfils its purpose as intended, but also that it supports the “act of improvement” better than other approaches. This statement is based on the test of three hypotheses, a regression analysis and further inspection of the experimental results. Two of our three assumptions are supported at a statistically significant level and the experiment confirmed that the BPI-Pattern Approach has a significant positive influence on both improvement effectiveness and efficiency. This makes the BPI-Pattern Approach a valuable asset when seeking for the improvement of existing business processes.

The contribution of this research is twofold and provides valuable insights for theory and practice. From a scientific point of view, it provides evidence that the newly developed artefact, the BPI-Pattern Approach, works as intended. We were able to confirm the common assumption that a more structured approach for BPI would generate better results (cf. [20], [24]) which could be explicitly demonstrated for the experimental scenario. Furthermore, the experiment helped to verify some of the qualities which are ascribed to patterns in general, such as providing guidance to novices [42] or reducing the effort of developing solutions [23], [24]. The contribution for practitioners is that the practical benefit of the BPI-Pattern Approach has been experimentally verified. Hence, it can be considered as a new instrument which expands the existing toolbox of methods, techniques and approaches that are used in improvement initiatives.

However, we still see some limitations which require further research regarding the evaluation of the BPI-Pattern Approach. There possibly exist other factors that have not been taken into consideration but may influence the effects of the pattern-based approach, and therefore should be part of future studies. These include but are not limited to e.g. usability aspects of the used software tool, the level and the complexity of the business process that has to be improved, the impact of group dynamics and teamwork etc. It would also be interesting to repeat the experiment with experienced BPI experts and to determine whether there are any differences regarding the approaches' benefits compared to our rather unexperienced user group. A long-term study could be used to validate if the advantages of the approach are stable over time or decrease e.g. because users have memorized the patterns and gained expertise.

The positive results of the evaluation refer to the particular setting which includes a number of simplifications due to the experimental design (e.g. number and size of the processes to be improved, number of available BPI-Patterns in the catalog etc.). Although the exemplary scenario was based on real-world data, the use of other evaluation methods, such as field experiments or case studies, which are able to capture the manifold interdependencies occurring in practice, is recommended to validate the findings in the context of real-world applications.

References

1. Harmon, P., Wolf, C.: The State of Business Process Management 2016. (2016)
2. Harrington, H.J.: Business process improvement : the breakthrough strategy for total quality, productivity, and competitiveness. McGraw-Hill, New York (1991)
3. Hammer, M., Champy, J.: Reengineering the corporation: a manifesto for business revolution. Harper Business, New York (USA) (1993)
4. Pande, P.S., Neuman, R.P., Cavanagh, R.R.: The Six Sigma way : how GE, Motorola, and other top companies are honing their performance. McGraw-Hill, New York (2000)
5. Kettinger, W.J.T., James T.C., Guha, S.: Business Process Change. A Study of Methodologies, Techniques, and Tools. MIS Quarterly March 1997, 55-80 (1997)
6. Siha, S.M., Saad, G.H.: Business process improvement: empirical assessment and extensions. Business Process Management Journal 14(6), 778-802 (2008)

7. Zellner, G.: A structured evaluation of business process improvement approaches. Business Process Management Journal 17(2), 203-237 (2011)
8. Nwabueze, U.: Process improvement: the case of a drugs manufacturing company. Business Process Management Journal 18(4), 576-584 (2012)
9. Falk, T., Griesberger, P., Johannsen, F., Leist, S.: Patterns for Business Process Improvement - A First Approach. 21st European Conference on Information Systems (ECIS 2013). Utrecht (The Netherlands) (2013)
10. Falk, T., Griesberger, P., Leist, S.: Patterns as an Artifact for Business Process Improvement - Insights from a Case Study. 8th International Conference on Design Science Research in Information Systems and Technology (DESRIST). Helsinki (Finland) (2013)
11. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28(1), 31 (2004)
12. Peffers, K., Tuunanen, T., Rothenberger, M., Chatterjee, S.: A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24(3), 45-77 (2007)
13. Rosemann, M.: Business Process Lifecycle Management. (2004)
14. Bhatt, G., D., Troutt, M., D.: Examining the relationship between business process improvement initiatives, information systems integration and customer focus: an empirical study. Business Process Management Journal 11(5), 532-558 (2005)
15. Smart, P.A., Maddern, H., Maull, R.S.: Understanding Business Process Management: Implications for Theory and Practice. British Journal of Management 20(4), 491-507 (2009)
16. Davenport, T., H.: Process Innovation - Reengineering work through Information Technology. Harvard Business School Press, Boston (USA) (1993)
17. Pourshahid, A., Mussbacher, G., Amyot, D., Weiss, M.: An aspect-oriented framework for Business Process Improvement. 4th International Conference MCETECH 2009. Ottawa, Canada: Springer (2009)
18. Griesberger, P., Leist, S., Zellner, G.: Analysis of Techniques for Business Process Improvement. 19th European Conference on Information Systems (ECIS 2011). Helsinki (Finland) (2011)
19. Valiris, G., Glykas, M.: Critical review of existing BPR methodologies. Business Process Management Journal 5(1), 65-86 (1999)
20. Povey, B.: The development of a best practice business process improvement methodology. Benchmarking for Quality Management & Technology 5(1), 27-44 (1998)
21. Gamma, E., Helm, R., Johnson, R., E., Vlissides, J.: Design Patterns - Elements of Reusable Object-Oriented Software. Addison Wesley (1995)
22. van der Aalst, W.M.P., ter Hofstede, A., H., M., Kiepuszewski, B., Barros, A., P.: Workflow Patterns. Distributed and Parallel Databases 14(3), 5-51 (2003)
23. Reijers, H.A., Limam Mansar, S.: Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics. Omega 33(4), 283-306 (2005)
24. Forster, F.: The Idea behind Business Process Improvement: Toward a Business Process Improvement Pattern Framework. BPTrends, 1-13 (2006)
25. Kim, D., Kim, M., Kim, H.: Dynamic Business Process Management based on Process Change Patterns. 2007 International Conference on Convergence Information Technology. Gyeongju (Korea) (2007)

26. Appleton, B.: Patterns for Conducting Process Improvement. Pattern Languages of Program Design 3 (PLoP Conference). Monticello (USA): Addison Wesley (1997)
27. Manns, M., L., Rising, L.: Fearless Change - Patterns for Introducing New Ideas. Addison Wesley (2005)
28. Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P., Stal, M.: Pattern-Oriented Software Architecture: A System of Patterns. John Wiley&Sons, Chichester (1996)
29. Paludo, M., Burnett, R., Jamhour, E.: Patterns Leveraging Analysis Reuse of Business Processes Software Reuse: Advances in Software Reusability. Wien (Österreich): Springer (2000)
30. Zellner, G.: Towards a framework for identifying business process redesign patterns. Business Process Management Journal 19(4), 600-623 (2013)
31. Lang, M., Wehner, B., Falk, T., Griesberger, P., Leist, S.: Evaluating Business Process Improvement Patterns by Simulation. 23rd European Conference on Information Systems (ECIS). Münster, Germany (2015)
32. March, S.T., Smith, G.F.: Design and natural science research on information technology. Decision Support Systems 15(4), 251-266 (1995)
33. Peffers, K., Rothenberger, M., Tuunanen, T., Vaezi, R.: Design Science Research Evaluation. 7th International Conference on Design Science Research in Information Systems and Technology (DESRIST 2012). Las Vegas (USA): Springer (2012)
34. Venable, J., Pries-Heje, J., Baskerville, R.: A Comprehensive Framework for Evaluation in Design Science Research. 7th International Conference on Design Science Research in Information Systems and Technology (DESRIST 2012). Las Vegas (USA): Springer (2012)
35. Cleven, A., Gubler, P., Huener, K., M.: Design Alternatives for the Evaluation of Design Science Research Artifacts. 4th International Conference on Design Science Research in Information Systems and Technology (DESRIST 2009). (2009)
36. Griesberger, P.: Developing the Evaluation of a Pattern-Based Approach for Business Process Improvement. 9th Conference on Design Science Research in Information Systems and Technologies (DESRIST 2014). Miami (USA): Springer International Publishing Switzerland (2014)
37. Mettler, T., Eurich, M., Winter, R.: On the Use of Experiments in Design Science Research: A Proposition of an Evaluation Framework. Communications of the Association for Information Systems 34, 223-240 (2014)
38. Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A.: Experimentation in Software Engineering. Springer, Heidelberg (2012)
39. Shadish, W.R., Cook, T.D.: Experimental and Quasi-Experimental Designs for generalized causal inference. Houghton Mifflin Company (2002)
40. Barros, O.: Business process patterns and frameworks. Business Process Management Journal 13(1), 47-69 (2007)
41. Rohleder, T., R., Silver, E., A.: A tutorial on business process improvement. Journal of Operations Managament 15(2), 139-154 (1997)
42. Jung, J., Sprenger, J.: Muster für die Geschäftsprozessmodellierung. DW 2006: Integration, Informationslogistik und Architektur. Friedrichshafen: Gesellschaft für Informatik (2006)

Systemizing Colour for Conceptual Modeling

Jeannette Stark, Richard Braun and Werner Esswein

Technische Universität Dresden, Germany[1]
(Jeannette.stark, Richard.braun, werner.esswein)@tu-dresden.de

Abstract. Colour is used in many conceptual models and is discussed intensively since MOODY has published his 'Physics of Notation'. Yet, choosing the right colour for a construct is difficult but crucial. Using a colour for a certain construct which is not appropriate can lead to visual stress as well as too much or too little emphasis on that construct. The aim of this paper is to give a systematization of colour for conceptual modeling by reviewing theories of colour vision, colour harmony and visual attention. Based on this review we provide colour combinations for different conceptual modeling colour scenarios.

Keywords: Colour, Hue, Conceptual Modeling, Pop-out, Perceptual Discriminability.

1 Introduction

Conceptual modeling constructs are mainly distinguished on the basis of the shape [1]. In BPMN 2.0, for example, shape (and to a small extent also texture) is used to derive the majority of constructs [2], which is useful, as shape has a significant impact on object recognition [3]. Yet, shape alone does not produce a high visual distance between modeling constructs and thus, leads to a low perceptual discriminability as only one visual variable is used to encode visual information. This makes it hard for model users to distinguish between constructs [1]. Redundantly using visual variables can help to increase perceptual discriminability and allows positive effects such as a faster detection of modeling constructs [1]. When visual variables are discussed for redundant coding, in most cases also the visual variable colour is included ([2],[4-6]). Colour is very powerful as differences in colour can be detected three time faster than shape and are more easily remembered [7-8]. Also experiments from visual attention confirm that colour dominates other variables [9-10].

So far, colour has been discussed to visually distinguish matching operators in Business Process Models (BPMs) [6], to produce a pop-out of chunks and of model elements [11] as well as to encode further information [12]. Colour has also been tested empirically for conceptual modeling in [6] and [13], who have shown that if colour is used, significant positive effects on comprehension for novices can be achieved. Yet, the application of colour still remains subjective. In studies that use colour for

[1] This research is supported by the Graduate Academy of TU Dresden, financed by the excellence initiative of the German federation and the federal states.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Stark, J.; Braun, R.; Esswein, W. (2017): Systemizing Colour for Conceptual Modeling, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 256-270

conceptual modeling, colour is not selected systematically. [13] argues that assigning a certain colour to a model element is very difficult due interrelations between hue and brightness. Even for designers choosing the right colour combination is very difficult. They often look for inspiration from sources such as art or colour palettes [14] and that way, try to avoid disharmonious colour combinations but rather create attractive ones which according to [15] works better than unattractive ones. Yet, designers of modeling grammars do normally not wish to spend time for being inspired which is why a systematization of colour for conceptual modeling would be helpful for them.

In this study we address the gap of a missing colour systematization for conceptual modeling by elaborating scenario-specific colour combinations that modelers can use during the modeling process if their modeling tool allows assigning a specific colour, that tool-designers can implement in their tools and that designers of modeling grammars can use when assigning colours to their constructs. We do not draw requirements for colour combinations from every possible perspective but limit our study to theories of Visual Attention (to derive scenario-dependent colours), of Colour Vision (to derive general and scenario-specific requirements) and of Colour Harmony (as in most cases several colours for several constructs are selected that should harmonize with each other). We do not discuss cultural-dependent colour perception as is recently discussed for conceptual modeling in [16] and remark that colour combinations of this paper are rather elaborated for a western context. We further do not treat semantics of colours which are also very cultural dependent.

This research uses design science [17] as research method. Following the research method discussed in [18] the paper is structured as follow: We start with a first initiative by reviewing conceptual modelling colour scenarios in section 2. Based on the colour scenarios we continue with a requirements analysis from theories of Visual Attention (section 3.1), Colour Vision (section 3.2) and Colour Harmony (section 3.3). Colour combinations for scenarios are presented as artefact and are implemented by giving an example in section 4. So far we have not evaluated the colour combinations but will present an idea of how this can be done in section 5, which further provides a short conclusion and an outlook of future research possibilities.

2 Colour in Conceptual Modeling

Colour has been used for different purposes in conceptual modeling: Production of a pop-out, a high visual distance as well as to encode further information (see table 1). [11] and [13] have used 3-4 colours to **produce a pop-out** of certain modelling elements and to visually distinguish further elements. While [11] made entity-types pop out from the rest of ERD-constructs, in [13] a pop-out effect is used for two chunks of BPMs. Also in [6] a pop-out of matching operators is discussed. Yet, in their model a great number of operators are visually distinguished by colour, which is why we argue that they rather aim to **produce a high visual distance** between operator pairs, as a pop-out can only be achieved for very few elements [13]. Further studies that use colour to produce a high visual distance for modeling constructs are [2], [19-21]. Colour is also used to **encode further information** in a non-redundant way in [12],[22-24]. The number of colours in these studies varies to 2-9 different colours.

Table 1. Colour application scenarios in conceptual modeling

Purpose of Colour Application	*# of colours*	*Mod. Gram.*	*Lit.*
Scenario 1: Pop-out of model elements and high visual distance of further objects			
Pop-out for important constructs: Pop-out for Entity-Types + high visual distance for Relationship-Types and Attributes	3	ERD	[11]
Pop-out for important chunks: Pop-out of parallel and alternative chunk + high visual distance for further chunks	4-6	BPM	[13]
Scenario 2: High visual distance			
High visual distance of *varying numbers* of operator *pairs:*	varying	BPM	[6]
High visual distance of a *fix number* of constructs:			
e.g. Gateways, DataObjects, Events	-	BPMN	[2]
e.g. Process, object, Team	9	UCM	[19]
e.g. Container, Link, Container Unit	2-9	WebML	[20]
e.g. Stages, WorkUnitKinds	4-8	SEMDM	[21]
Scenario 3: Encoding new information (non redundant application of colour)			
Start and Endpoint, nodes that require user's interaction	4	BPM	[22]
Debit and Credit information in places	4	Petri Nets	[23]
States of activities and delay indicator	5	BPMN	[12]
Coloured tokens and states	-	Petri Nets	[24]

3 Requirements for the scenarios

Requirements in this section are only discussed for scenario 1 and 2 as in conceptual modeling it is already consensus that colour should not be used in a non-redundant way [1-2]. In the next subsection requirements are derived from Visual Attention.

3.1 Requirements from Visual Attention

Colour is an effective variable to guide human attention [25], and can be used to produce a pop-out for the most important constructs of a model. A pop-out effect (as seen for the red entity types of fig. 1) is reached when visual variables are used in a way that parallel processing is facilitated [26]. In this case, visual information is perceived very fast, in less than 200ms, which allows working memory to efficiently process information. Yet, as shown by attention researchers such as [27], parallel processing is not a rigid state but a pole on a continuum, having serial processing as the other pole. Conditions for parallel and serial processing are discussed by [11] for conceptual modelling (see fig. 1). Colour can affect parallel or serial processing based on the number of unique values, its rang in feature hierarchy, its difference in value intensity and its similarity to other colours.

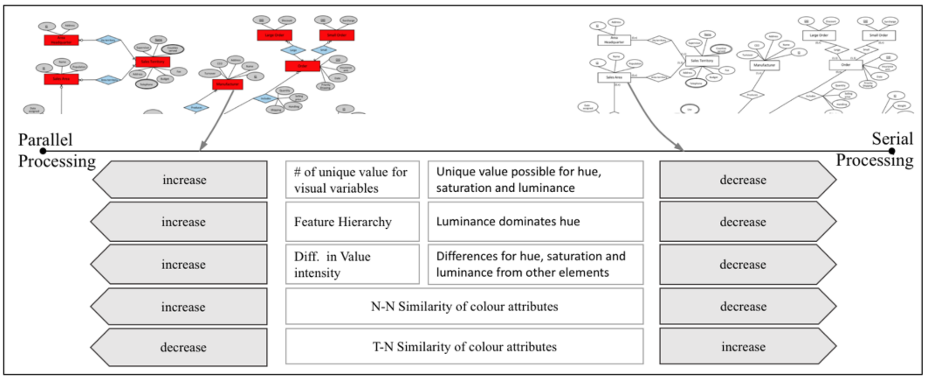


Figure 1. Continuum between parallel and serial processing on the basis of [11]

An element that is characterized with a unique value, such as circular for shape (fig. 2b), is in most cases further placed on the parallel side than an element with no unique values. That way, the circle in b) attracts more attention than the red circle in a), that does only have values which also appear in distractor elements. According to the theory of **Feature hierarchy** the visual system appears to have a hierarchy for visual variables (e. g. luminance dominates hue [25] and hue dominates shape [9]). That way, a unique value for hue (as in c) would lead to a more parallel place than a unique value for shape (as in b). A further condition concerns the **numbers of unique values**. In d) luminance is introduced as a further unique value besides hue (as in c). That way, the target element in d) is characterized with more unique values (hue und luminance) which leads to a more parallel position than the target element has in c). Working on the condition **differences in value intensity** does further help to move the element further to the parallel pole. In e) we have further increased the luminance difference between target and distractor elements and have, that way, increased parallel processing. When discussing differences in value intensity between target- and distractor elements, we implicitly talk about similarity (target-non-target similarity or T-N similarity). Besides T-N similarity also similarity between distractor elements matters (non target-non target similarity or N-N similarity) [27]. A target element is placed more on the parallel side with distractor elements having a high similarity (as is the case b-d) instead of having a lower similarity (such as in f).

For the two scenarios a different position of the modeling constructs along the continuum is required. For the pop-out scenario [13] discusses to place the constructs that should pop-out as far to the parallel side as possible (see fig. 3). Further constructs can be distinguished by giving them a position which is close to the serial processing side to avoid interferences with the pop-out constructs. For the high visual distance scenario all relevant constructs can be placed as much to the parallel processing side as possible. Based on the different positions within the scenarios the two scenarios require a different work on the conditions which are summarized in fig. 3 as requirements from Visual Attention.

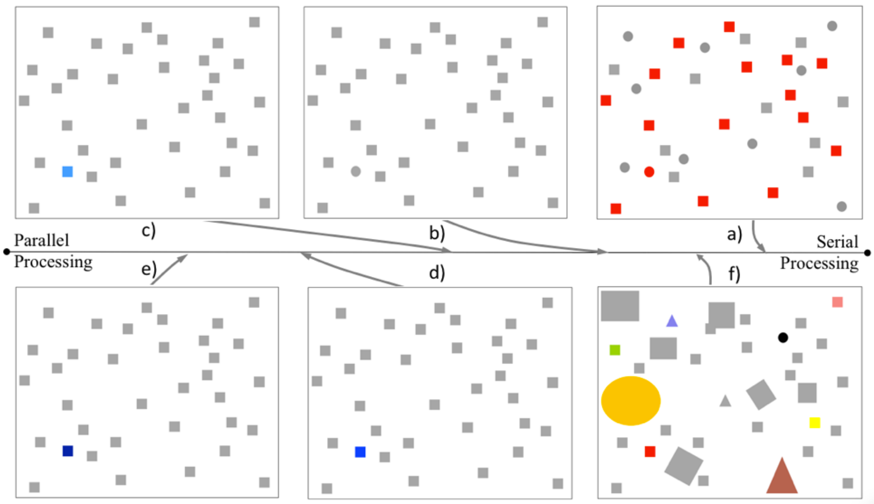
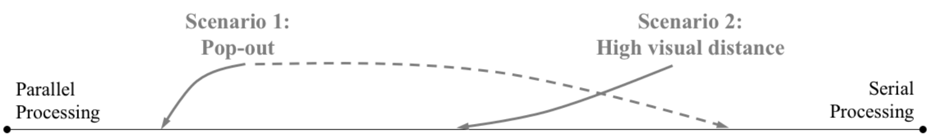


Figure 2. Examples to place elements along the continuum

For the pop-out scenario placing the most important constructs close to the parallel processing side can be achieved by creating a unique value for luminance which is highest in feature hierarchy and holding this value constant among other constructs. That way, a clear focus is achieved with the most important variable of the feature hierarchy (req. 1 & 2). Furthermore, a greater number of unique values should be created with hue and saturation (req. 3) and value differences should be set as high as possible (req. 4). If possible, sufficient differences in value intensity of non pop-out constructs should be regarded, too, to further visually distinguish these constructs (req. 5). For the high visual distance scenario, a pop-out is not required for the constructs but visual distances should be arranged as high as possible which requires unique values for those constructs with sufficient luminance and hue contrast (req. 1 & 2).



Req. 1: Use sufficient luminance contrast between the pop-out colours and non pop-out colours

Req. 2: Use constant luminance for non pop-out colours

Req. 3: Produce futher unique values for pop-out colours

Req. 4: Sufficient difference in value intensity between pop-out and non pop-out colours

Req. 5: Sufficient difference in value intensity for attributes of non-pop-out colours to further distinquish non pop-out elements

a)

Req. 1: Use sufficient luminance contrast between the colours

Req. 2: Use sufficient hue contrast between the colours

b)

Figure 3. a) Scenario-related positions along the continuum and b) requirements from Visual Attention

3.2 Requirements from Colour Vision

Implications from colour vision have already been discussed within 38 guideline for computer graphics in [28]. In this work we discuss 13 of the more general guidelines of [28] which can also be applied for conceptual models (on papers and on screens) and whose focus fits to our domain.

Human colour vision is trichromatic [29]. The retina has three types of colour-sensitive photoreceptors that are referred to as S, M, and L cones, which are sensitive to **s**hort- , **m**iddle- as well as **l**ong-wavelength of the spectrum [30] and are usually termed blue, green and red. We perceive colour by combining the three receptor inputs into a unique triplet and thus, perceive colour depending on the ratio of the three cone-responses. S, M, and L-cones are not equally distributed among the retina but occur in relation of 1:20:40, which means that the sensitivity to red and green is much higher than to blue [31]. A difference in contrast sensitivity has also been measured for the neural channels that distribute colour information to the brain. In this process S, M and L-wavelength are transformed into an achromatic and two chromatic signals [32]. While the achromatic signal represents the sum of the cone responses, the two chromatic signals define colour differences in red-green and yellow-blue. These channels differ in contrast sensitivity, too. The achromatic channel offers highest contrast sensitivity and the yellow-blue channel the lowest, which can be explained with the relative distribution of cones in the central fovea [28].

In consistency with the chromatic and achromatic channels, HERING proposed his **opponent colour theory** of colour vision around 100 years ago [28]. He noted that certain hues never occur together such as red-green or yellow-blue in the sense that a colour is not described as reddish-green. Based on these observations HERING argued that there is something fundamental about those pairs of colours [30] and used these colours together with white and black to define a system of six elementary colours (see fig. 4a). Out of this system three perceptual dimensions can be derived to describe colour [30]: Hue, colourfulness and brightness. While hue defines if an area appears to be similar to red, green, yellow and blue or to a combination of those colours, colourfulness describes to what extend an area appears chromatic. Brightness gives information about the extend an area emits of reflects light [33].

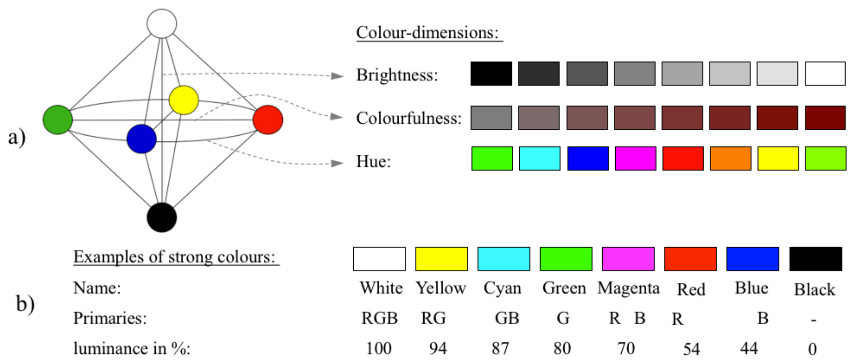
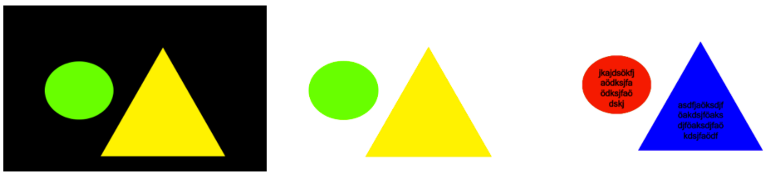


Figure 4. a) Colour dimensions, and b) examples of strong colours (on the basis of [28])

Further terms that are used in this paper are saturation (which defines colourfulness in relation to the brightness of similarly illuminated white [33]), and luminance (which determines brightness sensitivity of the human eye [34]). While green, red and blue represent primary colours that can directly be detected by S-, M-, and L-cones, secondary colours present a ratio of the primary colours [28]. In fig. 4b) saturated primary and secondary colours are ordered according to their luminance. These colours with luminances close to that of fig. 4b) are further referred to as strong colours. According to [28] bright **strong colours** can be used to claim for attention [28], which is also the aim of pop-out colours (see fig. 5a). Yet, these colours are discussed for a black background which allows sufficient luminance differences. In conceptual modeling a black background is not common, which is why strong colours should rather contrast to a white background (which is not achieved in b). Fig. 4b describes strong colours which are helpful to select a pop-out colour that offers enough contrast to the background. While a white background has a luminance of 100% a sufficient contrast is offered in saturated red (54%) or saturated magenta (70%). Saturated blue also offers a high difference, but if text is used inside the constructs, it cannot be read (c).



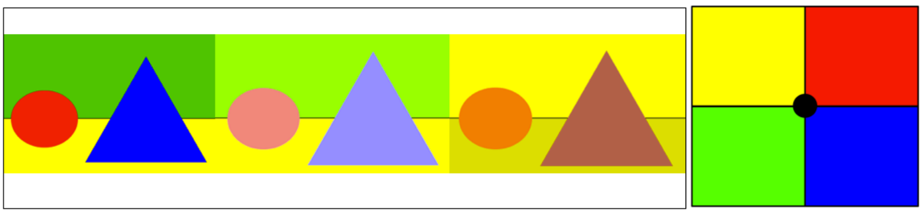
a) b) c)

Figure 5. a) Las Vegas effect: bright, strong colours on a black background, b) the same colours on a white background and, c) colours with a sufficient luminance contrast.

A certain colour appearance of an element does not only depend on the element's own colour attributes but also of colour attributes of neighboring regions or elements [28]. According to **simultaneous contrast**, colours shift in appearance when the colour of the background changes. These shifts follow HERING's opponent theory of colour vision. That way, a light background induces a stimuli to appear darker and smaller and a dark background induces a lighter and larger appearance. Furthermore, red induces green, green induces red, yellow induces blue, and blue induces yellow [30]. That is why [28] suggests to use a neutral mid-gray background, whenever accurate visual colour-judgment is required. Yet, since in conceptual modeling white has always been used as background and an accurate judgment of colours is in most cases not required, we have decided not to use the grey background as a colour guideline.

Looking at strong colours can fixate a coloured image on the retina, which temporarily reduces the sensitivity of photoreceptors and leads to **afterimages** [28]. This can be experienced when fixating the gaze on the black dot of fig. 6d) for at least 10s. When looking on a white paper afterwards an afterimage with the same layout but complementary colours can be observed. That way, in formerly red regions of the image cyan emerges because the response of long-wavelength (red) is temporarily suppressed which leads only the medium- (green) and short- (blue) wavelength to emerge [28]. This is why generally it is not recommended to use large areas of bright colour [35].

Strong colours can also lead to a **depth contrast**. When light is passing through the lens it is refracted into its spectral components. Depending on the wavelength, spectral components converge beyond (red) or in front of the retinal surface (green and blue). That way, to reach a focus on red colours the lens have to become more convex while a focus on blue or green colours requires the lens to become less convex like seeing an object which is far away. This is why, for most observers red appears to advance and blue to recede [36] which can lead to unwanted depth effect when putting strong red and strong blue on neighboring areas. Yet, this effect also can be used to impose colour contrast for a pop-out effect (see fig. 5c). While strong colours are very useful as pop-out colours, these colours should be used very carefully for the high visual distance scenario. In this case we need to place the constructs as far as possible on the parallel processing side without producing visual stress (which we accept for the pop-out scenario when moving the constructs even further to the parallel side). A strong colour combination used for large areas such as in fig. 6a) imposes visual stress due to too much contrast. We can avoid visual stress by using two strategies: First, using colours with a high luminances above the level described in fig. 4b does lead to a colour which is not strong and which would (due to the high luminance-level) not produce too much contrast to a white background (b). When using this strategy, colours should still have enough luminance differences as luminance is highest in feature hierarchy and is thus, important to discriminate between elements. Second, using low hue differences also leads to reduced contrast, which in turn reduces visual stress (c). Moreover, the size of the constructs matters. As visual stress occurs when using strong colours for large areas we should consider that colours for smaller areas need also to be stronger.



a) b) c) d)

Figure 6. a-c: Visual stress within the high visual distance scenario, and d) example for the afterimage effect

If **text** is required within a coloured object then legibility becomes important, which is usually facilitated with a sufficient luminance contrast ratio between text and background. If coloured text on coloured background cannot be avoided the luminance contrast ratio should be at least 3:1 and preferably 10:1 [28].

Human observers with nondeficient colour vision depart to some extend from the standard spectral responses which is why human observers do not see a given colour in exactly the same way [28]. **Colour blindness** further leads model users not to perceive colour the way it is intended it to be seen. Approximately every 12. male and every 100. female adult in Europe is colour deficient [37]. In most cases the kind of colour-deficiency is dichromatism, where one of the three cone pigments is missing.

Dichromats have problems discriminating hues [38], which is why it is recommended not to use hue alone to discriminate between elements. A further restriction of colour concerns the **number of colours** that should be used. Two factors restrict the number of colours [28]: The ability to simultaneously process a certain number of colours and the ability to discriminate between different colours. MILLER has estimated that humans can simultaneously process 7 +/-2 elements in working memory [39]. Recent research estimates this amount lower to 3-4 items [40]. According to research in psychophysics we can differentiate around 7-10 colours [41-42]. Based on those limitations a general limitation of 5-7 colours is given by [43], which seems reasonable based on the number of colours we can differentiate and process. Implications of the description of Colour Vision for conceptual modeling are summarized in table 2.

Table 2. General guidelines for colour application (based on [28])

1.	Be aware that strong colours in neighbouring and in big areas can produce visual stress from unwanted depth effects and afterimages.
2.	Do not use the blue-yellow channel alone for fine details.
3.	Do not use hue alone to encode information
4.	Be aware that colours of neighbouring areas change the appearance of colours.
5.	Do not use coloured text on coloured background.
6.	Limit the numbers of colours to 5-7.

Based on these general guidelines also implications for the pop-out scenario as well as for the high visual distance scenario have been discussed and are summarized in fig. 7.

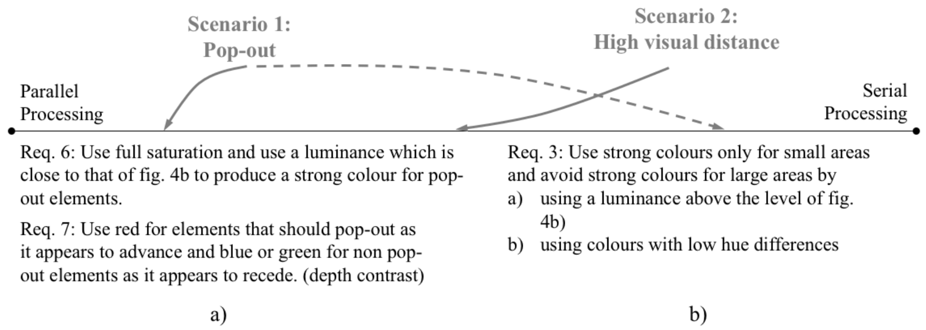


Figure 7. Scenario-specific requirements from colour vision

3.3 Requirements for an effective combination of colours

When applying the rules of section 3.1 and 3.2 there is still a lot of subjectivity in choosing the right colour combination for model elements. Yet, choosing harmonious colours is difficult and crucial [14]. Especially, when tasks need to be solved under stress a visual graphic, used for this tasks, works better if it is designed aesthetically and harmoniously [15]. Choosing the right colours has been challenged for already a long time starting with the ancient Greeks. Modern colour theory has began with the colour wheel that has been proposed by NEWTON and was further developed by

GOETHE. Later, hue templates on colour wheels were developed [44]. MATSUDA, for example, created a set of eight hue templates that describe areas on a colour wheel in which harmonious colours can be selected. Hue templates can easily be rotated around the colour wheel which leads different harmonic hue combinations to appear and are often used as a starting point by designers [14]. A plethora of colour theories address colour harmony beyond just hue. Many of these theories suggest that colours are harmonious if one colour attribute contrasts while the others remain fixed [14]. That way, the Munsell system suggests colour sets being fixed in value and hue but vary in saturation to be harmonious [45] while the Ostwald system proposes colours with equal white or black content as harmonious [46]. Those suggestions were only tested in recent years (see for example [47]) and have led to contradictory results although a few trends have been found [14]: Colours harmonize if they are characterized with same hue, equal or similar saturation as well as contrasting lightness value. In recent years a broad empirical basis through social colour networks, such as Adobe Kuhler [48] and COLOURLovers [49], has emerged that offer millions of colour palettes and have been used by colour scientists to test their theories. Both social colour networks allow their users to create colour themes consisting of 1-5 colours as well as rating, commenting and modifying them. Based on the contradictory results from theories of colour harmony we decided to use colour palettes from two social colour networks that have also been used in colour studies (Adobe Kuhler as well as COLOURLovers).

4 Colour combinations to create a pop-out and a high visual distance

To exemplarily show how colour combinations can be used for the different colour scenarios, we have searched colour palettes that have at least 100 “likes” or “loves” in COLOURLovers or Adobe Kuhler. These colour palettes have further been assessed according to the requirements of section 3 for scenario 1 (pop-out) and 2 (high visual distance). Those colour palettes that best fulfilled the requirements were chosen as colour combination for conceptual modeling. Fig. 8 and 9 give an overview of most appropriate colour combinations for the scenarios and shows to which extend these combinations fulfill the requirements. For the pop-out scenario we found 11 and for the high visual distance scenario we have extracted 22 colour palettes. The colour combinations for the pop-out scenario can be found on the left side of fig. 8 and are evaluated according to specific pop-out requirements as well as legibility as a general requirement of section 3 on the right side of fig. 8. We found two colour combinations that provide two pop-out colours (on the top of fig. 8) and nine colour combinations with only one pop-out colour (on the bottom of fig. 8). The remaining colours within these colour combinations can be used to visually distinguish elements that should not pop-out and are ordered from smaller to larger elements. Within the separation of one-pop-out colour combinations and two-pop-out colour combinations the colours are ordered according to how they fulfil the requirements of the section 3.

Scenario 1: Pop-out

Legend:
pop-out colour | colour to produce visual distance

X_z: criterion is fulfilled
O_z: criterion is partly fulfilled
z: min. luminance contrast in % in req. 1
max. luminance contrast in % in req. 2

					Use sufficient luminance contrast between pop-out and non pop-out colours (req.1)	Use constant luminance for non pop-out elements (req. 2)	Produce further unique values for pop-out colours (saturation and hue) (req. 3)	Pop-out colour is saturated (req. 6)	Use red hues for pop-out element and green or blue hues for non pop-out elements (Req. 7)	Sufficient difference in hue and/or saturation between non pop-out elementns (Req. 4)	Further discrimination between non pop-out elements is possible (Req. 5)	Readibility of black-coloured text
E8608C	71CBC4	CDD56E	FFBD68	FFF9F4	O_7	X	X	-	O	X	-	X
DB3026	7ABF66	F9E14B	EFED89	E88A25	X_{16}	O_3	X	-	O	O	-	X
FF2121	C2FC63	BCF7EF	D7EEFA	FD9A42	X_{33}	O_4	X	X	X	X	X	X
FF634D	FDEDD0	FFF0AA	BCF1ED	FD795B	X_{28}	O_3	X	X	-	X	X	X
FF3F7F	F1CC5D	D6DD54	ADDFE3	F1DC9D	X_{18}	O_3	X	X	-	X	X	X
FF2121	BCF7EF	D7EEFA	C2FC63	FD9A42	X_{36}	X	X	X	X	X	X	X
FB0C06	D7EDA	CEECEF	FFC52C	030D4F	X_{37}	O_2	X	X	X	-	O	X
CC0C39	C8CF02	95CFB7	F8FCA7	E6781E	X_{30}	X	X	.	X	X	O	X
FA2A00	D6D8A8	F2D694	86B8B1	3D1C00	X_{28}	O_2	X	X	-	-	O	X
AE2F27	85B394	A7BA59	F9F0D8	F9D890	X_{25}	O_2	X	-	X	O	O	O
B42310	B0E629	F7CF0A	FCE70D	FA7C07	X_{40}	O_2	X	-	-	X	X	O

Figure 8. Pop-out colour combinations

The colour combinations for producing a high visual distance are summarized in fig. 9 and are further distinguished in colour combinations for large (top of fig. 9), small (in the middle) and a combination of large and small areas (on the bottom). The colour combinations are further ordered based on how many colours they offer and on how they fulfill the requirements of section 3. The colour combinations on the bottom of fig. 9 for large- and small-area colours are further distinguished based on the application for large- and small-areas. That way, the small-area colours are presented on a gray background while the large-area colours are presented on a white background. In most cases not every colour of the palette fits into the scenario which is why these colours are depicted in black in fig. 8 and 9. For a better understanding of fig. 8 and 9 we exemplarily show how these colour combinations can be used for conceptual modeling. Therefore, we discussed the lifecycle diagram discussed in [21], p. 200 using the graphical notation of ISO/IEC 247744. This lifecycle diagram uses colour to visually distinguish the modeling constructs (scenario 2). In its original version on the left side of fig. 10a) this diagram imposes visual stress by using eight different colours (4 line and 4 background colours for the constructs). By limiting colours of the resulting model (on the right side of fig. 10a) to the background colours we have reduced visual stress. The original background colours further fulfils the requirement 3a) by using colours that are above the luminance level of strong colours.

Scenario 2:

Large-area colours

X_z: criterion is fulfilled O_z: criterion is partly fulfilled

z: min. luminance diff. (%) in req. 4 and hue distance in degree in req 5

						Readability of black-coloured text	No use of strong colours (req. 3)	High luminance or low hue differences (req. 3a&b)	Sufficient luminance contrast between colours (req. 1)	Sufficient hue contrast (req .2)
5 -colours	D9B2FF	BAF3C3	F9FC9D	FFCBCF	DDE5FE	X	X	X	O_2	X_{45}
	D9ABFF	ABE4FF	FFABAB	DDFFAB	FFDAAB	X	X	X	O_1	O_{34}
	A7D3D2	D3DBB2	FFC6BC	FEDEA2	FFF9B8	X	X	X	O_1	O_{16}
	FFA398	9AD9D2	D0F7A6	FFC48C	FCE5C0	X	X	X	O_1	O_8
	F3BA2B	A2C5D6	D6DB89	FFD876	F6EFCD	X	X	X	-	-
4 colours	FFDAAB	DDFFAB	ABE4FF	D9ABFF	FFABAB	X	X	X	O_2	X_{50}
	93DFB8	FFC8BA	E3AAD6	B5D8EB	FFBDD8	X	X	X	O_1	X_{52}

Small-areas colours

X_z: criterion is fulfilled O_z: criterion is partly fulfilled

z: min. luminance diff. (%) in req. 3 and hue distance in degree in req 4

						Readability of black-coloured text	Use of strong colours (req. 3)	Sufficient luminance contrast between colours (req. 1)	Sufficient hue contrast (req .2)
5 colours	AAFF00	FFAA00	FF00AA	AA00FF	00AAFF	X	X	O_3	X_{40}
	FF0012	FFD900	5BE300	0084B0	FF7D00	O	X	X_5	O_{22}
4 colours	D60000	FFC801	93C700	0E99DA	FF530D	X	X	X_{10}	O_{29}
	98C000	EA2E49	FFE11A	0CDBE8	3D4C53	X	X	X_{10}	O_{21}
	7B9EF8	EB5153	FBE230	A1E736	010101	X	X	X_9	O_{31}
	0FAAB1	EB265D	FF9326	B7E01F	332C1F	O	X	X_9	X_{43}
3 colours	FF007C	00FFFF	FFFF00	00FF00	FF5100	X	X	X_{14}	X_{60}
	54BFAC	F2E530	D94625	1BA68C	F2EDA7	X	X	X_{16}	X_{45}
	D80351	F5D908	00A3EE	929292	3F3F3F	X	X	X_{13}	X_{146}
	FFCD4A	E6503B	6AACB8	E5E4C7	FFFFEA	X	X	O_8	X_{36}

large and small-area colours

X_z: criterion is fulfilled O_z: criterion is partly fulfilled

z: min. luminance diff. (%) in req. 5 and hue distance in degree in req 6

Small-area colour Large-area colour

						Readability of black-coloured text	Large area colours: avoid strong colour (req.3)	Large area colours: high luminance or low hue differences (req. 3a&b)	Small-area colour: strong colour(req.3)	Sufficient luminance contrast between colours (req. 1)	Sufficient hue contrast (req .2)
5 -colours	E87657	98B4BF	FF9B6F	E8D495	FFF3D2	X	X	X	O	O_3	O_3
	ED7D7C	A3AC7D	D4C687	F2C778	F1DAA8	X	X	X	O	O_3	-
	F8F087	F39DD4	DBBAE5	B7E3C0	B8D0DD	X	X	X	O	-	X_{36}
4 colours	FE6960	FEBF97	D2FDFE	FEFAC2	AFFBFF	X	X	X	O	O_4	O_{20}
	45AAB8	F06B50	FAFABA	E1D772	394240	X	X	O	X	X	-

Figure 9. Colour combinations for producing a high visual distance

Yet, the combination of colours might not seem harmonious to every user. This is why we have chosen a harmonious colour combination from fig. 9b with higher luminances for the larger elements using FFF9B8 and FEDEA2 (ranging from 88 to 96%) and lower luminances for smaller elements using A7D3D2 and A7D3D2 (ranging from 79-83%) to produce a stronger contrast to the white background. For these colours the min. luminance difference is 4% and luminance is used in a range of 17%. Hue differences are not as high as in the original model. We have used a hue distance of at least 107° to other colours for the process kind, as this construct is probably the most important construct. The other hues only range about 33° and do that way not produce to much contrast, and hence visual stress.

Figure 10. a) example model discussed in [21] before (left) and after (right) using a harmonic colour combination, b) details of harmonic colour combination

5 Conclusion

With this paper we have discussed requirements from different fields of theories for two different conceptual modeling colour scenarios. These requirements were further used to assess harmonic colour combinations of two different social colour networks. Moreover, we have shown exemplarily how these colour combination can be used to assign harmonic colours to a conceptual model. We found that using harmonic colour combinations defined in social colour networks has advantages and disadvantages. On the one hand, by using only colours that have been rated as good we can make sure that the model that uses these colours appears harmonic which is important for problem solving [15]. On the other hand, requirements drawn in section 3 were often not fulfilled. That way, luminance differences between pop-out and non pop-out constructs as well as hue contrast between colours for producing a high visual distance can be optimized if colours are picked individually instead of using a colour palette. This trade-off of using colour-palettes still needs to be evaluated. The experiment described in [13] might offer a starting point (by using colours from fig. 8 in the experiment model) for an evaluation of effects that might result from using colour palettes for the pop-out scenario and might give further insights of the trade-off of using colour-palettes.

References

1. Moody, D.: The "physics" of notations: Toward a scientific basis for constructing visual notations in software engineering. Softw. Eng. 35, 756-779 (2009)
2. Genon, N., Heymans, P., Amyot, D.: Analysing the cognitive effectiveness of the BPMN 2.0 visual notation. In: Int. Conf. on Softw. Lang. Eng. pp. 377–396. Springer (2010).
3. Biederman, I.: Recognition-by-components: a theory of human image understanding. Psy. Rev. 94, 115 (1987)
4. Moody, D., van Hillegersberg, J.: Evaluating the visual syntax of UML: An analysis of the cognitive effectiveness of the UML family of diagrams. In: Softw. Lang. Eng. pp. 16 (2008)
5. Moody, D.L., Heymans, P., Matulevičius, R.: Visual syntax does matter: improving the cognitive effectiveness of the i* visual notation. Requir. Eng. 15, 141-175 (2010)
6. Reijers, H.A., Freytag, T., Mendling, J., Eckleder, A.: Syntax highlighting in business process models. Dec. Supp. Syst. 51, 339–349 (2011)
7. Lohse, G.L.: A cognitive model for understanding graphical perception. Hum.-Comp. Interact. 8, 353–388 (1993)
8. Treisman, A.: Perceptual grouping and attention in visual search for features and for objects. J. Exp. Psy. Hum. Perc. Perf. 8, 194 (1982)
9. Callaghan, T.C.: Interference and dominance in texture segregation: Hue, geometric form, and line orientation. Perc. Psychophys. 46, 299–311 (1989)
10. Snowden, R.: Texture segregation and visual search: A comparison of the effects of random variations along irr. dimensions. J. Exp. Psy. Hum. Perc. Perf. 24, 1354–1367 (1998).
11. Stark, J.: Perceptual Discriminability in Conceptual Modeling. In: Enterprise Engineering Working Conference. pp. 103–117. Springer (2016).
12. Momotko, M., Nowicki, B.: Visualisation of (distributed) process execution based on extended BPMN. In: Database and Expert Systems Applications, pp. 280–284 (2003).
13. Stark, J., Braun, R., Esswein, W.: Perceptually discriminating Chunks in Business Process Models. In: CBI, Paris, pp. 84-93 (2016)
14. O'Donovan, P., Agarwala, A., Hertzmann, A.: Color compatibility from large datasets. ACM Trans. Graph. TOG. 30, 63 (2011).
15. Norman, D.: Emotion & design: attractive things work better. Interactions. 9, 36–42 (2002)
16. Kummer, D.T.-F., Leimeister, P.D.J.M., Bick, P.D.M.: On the Importance of National Culture for the Design of Information Systems. Bus. Inf. Syst. Eng. 4, 317–330 (2012)
17. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Quarterly 28, 75–106 (2004)
18. Verschuren, P., Hartog, R.: Evaluation in Design-Oriented Research. Qual. Quant. 39, 733–762 (2005)
19. Genon, N., Amyot, D., Heymans, P.: Analysing the cognitive effectiveness of the UCM visual notation. In: Int. Works, on Syst. Anal. and Modeling. pp. 221–240. Springer (2010)
20. Granada, D., Vara, J.M., Brambilla, M., Bollati, V., Marcos, E.: Analysing the cognitive effectiveness of the webml visual notation. Softw. Syst. Model. 1–33 (2013)

21. Sousa, K., Vanderdonckt, J., Henderson-Sellers, B.: Evaluating a graphical notation for modelling software development methodologies. J. Vis. Lang. Comp. 23, 195–212 (2012)
22. Koschmider, A., Caporale, T., Fellmann, M., Lehner, J., Oberweis, A.: Business process modeling support by depictive and descriptive diagrams. IN: EMISA (2015)
23. Werner, M.: Colored petri nets for integrating the data perspective in process audits. In: Int. Conf. on Conceptual Modeling. pp. 387–394 (2013)
24. Jensen, K., Kristensen, L.M.: Colored Petri Nets: A Graphical Language for Formal Modeling and Validation of Concurrent Systems. Commun ACM. 58, 61–70 (2015)
25. Callaghan, T.C.: Dimensional interaction of hue and brightness in preattentive field segregation. Perc. Psychophys. 36, 25–34 (1984)
26. Healey, C.G., Enns, J.T.: Attention and visual memory in visualization and computer graphics. Vis. Comp. Graph. 18, 1170–1188 (2012)
27. Duncan, J., Humphreys, G.: Visual search and stimulus similarity. Psy. Rev. 96, 433 (1989)
28. MacDonald, L: Using color effectively in computer graphics. Comput. Graph. Appl. 19, 20–35 (1999)
29. Young, T.: On the Theory of Light and Colours, The Philosophical Transaction, 1801. Course Lect. Nat. Philos. Mech. Arts II. 613–632 (1801)
30. Fairchild, M.D.: Color appearance models. John Wiley & Sons (2013)
31. Hunt, R.W.G., Pointer, M.R.: Measuring colour. John Wiley & Sons (2011)
32. Wandell, B.A.: Foundations of vision. Sinauer Associates (1995)
33. ILV, C.: International lighting vocabulary. Vienna CIE. (2011)
34. ASTM. Standard Terminology of Appearance. E284-09a., (2009)
35. Murch, G.: Physiological principles for the effective use of color. Comp. Graph. Appl. 4, 48–55 (1984)
36. Thompson, P., May, K., Stone, R.: Chromostereopsis: a multicomponent depth effect? Displays 14, 227–234 (1993)
37. Birch, J.: Worldwide prevalence of red-green color deficiency. JOSA 29, 313–320 (2012)
38. Meyer, G.W., Greenberg, D.P.: Color-defective vision and computer graphics displays. Comput. Graph. Appl. 8, 28–40 (1988)
39. Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psy. Rev. 63, 81 (1956)
40. Cowan, N.: The magical mystery four how is working memory capacity limited, and why? Curr. Dir. Psy. Sci. 19, 51–57 (2010)
41. Bertin, J.: Semiology of graphics: diagrams, networks, maps. Univ. of Wisc. Press (1983)
42. Stevens, S.S.: Psychophysics. Trans. Publishers (1975)
43. ISO 9241, Ergonomic Requirements for Office Work with Visual Display Terminals, Part 12: "Presentation of Information. ISO, Geneva, Switherland (1998)
44. Matsuda, Y.: Color design. Asakura Shoten. 2, 10 (1995)
45. Munsell, A.H., Birren, F.: A grammar of color. Van Nostrand Reinhold (1969)
46. Ostwald, W.: Colour science. (1931)
47. Ou, L.-C., Luo, M.R., Woodcock, A., Wright, A.: A study of colour emotion and colour preference. Colour emotions for single colours. Color Res. Appl. 29, 232–240 (2004)
48. Adobe Color CC, https://color.adobe.com/de (assessed 25/19/2016)
49. COLOURlovers, http://www.colourlovers.com/ (assessed: 25/10/2016)

Towards Situational Reference Model Mining – Main Idea, Procedure Model & Case Study

Jana-Rebecca Rehse[1] and Peter Fettke[1]

[1] Institute of Information Systems at the DFKI and Saarland University,
Campus D3.2, 66123 Saarbrücken, Germany
{Jana-Rebecca.Rehse, Peter.Fettke}@iwi.dfki.de

Abstract. This contribution introduces the concept of Situational Reference Model Mining, i.e. the idea that automatically derived reference models, although based on the same input data, are intended for different use cases and thus have to meet different requirements. These requirements determine the reference model character and thus the technique that is best suited for mining it. Situational Reference Model Mining is based on well-known design principles for reference modeling, such as configuration, aggregation, specialization, instantiation, and analogy. We present a procedure model for Situational Reference Model Mining and demonstrate its usefulness by means of a case study. Existing techniques for Reference Model Mining are examined and mapped to their underlying design principles. This way, we are not only able to provide reference model designers with concrete guidelines regarding their choice of mining technique, but also point out research gaps for the development of new approaches to reference model mining.

Keywords: Reference Model Mining, Reference Model Design, Reference Model Design Principles, Reference Model Construction, Inductive Reference Model Development

1 Introduction

Reference models can be considered as special conceptual models that serve to be reused for the design of other conceptual models [1, 2]. They provide a template for process models in a certain industry and thus facilitate a resource-efficient implementation of the respective process and its adaption to the individual needs of an organization. This way, companies may benefit from best practices and industry-specific experience. The use of reference models is associated with a higher quality of processes and process models, as it simplifies internal communications by introducing a common terminology and considerably reduces the resources required for business process management [3].

Given a reuse-oriented conceptualization of reference models, their main purpose is to serve as an orientation in the design of new business process models. In this context, we decipher two general design processes [4]. Deriving an individual model from a reference model is known as "Design With Reuse" (DWR), i.e. an existing

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Rehse, J.-R.; Fettke, P. (2017): Towards Situational Reference Model Mining - Main Idea, Procedure Model & Case Study, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 271-285

model is used as a blueprint offering guidance to the process model designer by giving suggestions for both content and design of the individual model. On the other hand, "Design For Reuse" (DFR) describes the process of constructing a (reference) model for the purpose of being reused, i.e. composing model parts and domain knowledge, such that they achieve a certain degree of universality.

Considering a model construction process, there exist several different techniques for deriving a conceptual model from another one. These so-called design principles describe how the content of the original model is adopted, adapted, and extended in order to create a new model. Five design principles are described in the literature [2]. Each configuration, instantiation, specialization, aggregation, and analogy may be used in the context of reference modeling and applied to both DFR and DWR.

Not every design principle may be applied to every reference model, nor may every intended target model be derived by any design principle. The principles differ in terms of concretization and usability. For example, the Configurable EPC (C-EPC) constitutes an application of the configuration principle to Event-Driven Process Chains [5]. Instead, the choice of model design principle depends on the situational circumstances, i.e. the requirements posed to the target model and the construction process itself. These factors also determine the character and thus the choice of an appropriate reference model for a certain application context.

Considering the situational circumstances and requirements is especially important when designing a reference model inductively, i.e. deriving it (semi-)automatically from a set of individual process models (Reference Model Mining). Besides the input models, the reference model content and character is determined by the choice of the mining technique. Different mining techniques yield different models, even when applied to the same set of input models. As the desired outcome depends on the intended reuse of the mined model (Design With Reuse), the situational circumstances also determine the recommended, or preferable, mining technique (Design For Reuse). This concept is called "Situational Reference Model Mining (S-RMM)", i.e. extending RMM towards consciously considering the situational context when designing and using a reference model.

In this contribution, we follow a design-science research approach [6] in order to elaborate how existing concepts in reuse-oriented reference modeling can be applied to the relatively new field of Reference Model Mining. How to develop guidelines for reference model designers? What constitutes a context for applying S-RMM? Which concrete mining techniques instantiate which principle?

Therefore, we introduce important foundations in reuse-oriented reference modeling, reference model design principles, and reference model mining in Section 2 and analyze Related Work and the emerging research gaps in Section 3. Section 4 introduces the concept of S-RMM by explaining the conceptualization and idea, defining a procedure model, and analyzing existing mining techniques regarding their application in a situational context, in order to give concrete guidelines to reference model designers. In Section 5, the procedure model and accompanying guidelines are applied in terms of a case study. The paper is concluded with a discussion and an outlook in Section 6.

2 Foundations

2.1 Reuse-Oriented Reference Modeling

Reusing a reference model entails adopting the contents of a model as well as adapting and extending them to fit the specific application context. Figure 1 outlines a reuse-oriented reference model construction [2, 4]. In a typical model construction process, the model designer creates a model according to the user's requirements, employing specific methods. The construction process is influenced by both the model quality (effectiveness) and the required time and cost (efficiency). Reference models can be understood as tools that foster both the effectiveness and the efficiency of model construction. They include contents that are relevant for different application contexts ("Design For Reuse") and may serve as the basis for several construction processes ("Design With Reuse"). As the model contents do not have to be newly constructed and have already been applied, both effectiveness and efficiency are increased.

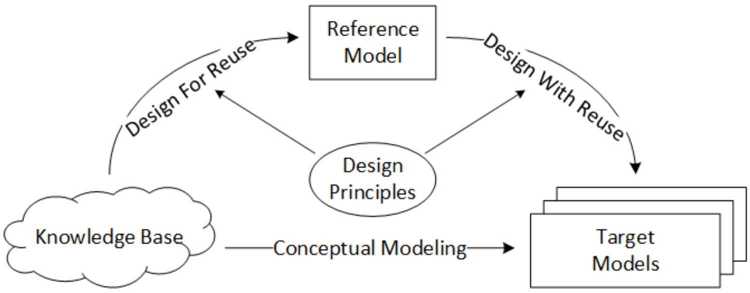


Figure 1: Reuse-Oriented Reference Modeling (cf. [2, 4])

2.2 Reference Model Design Principles

A design principle is a rule that describes how the content of one model is used in the construction process of another. This entails adopting as well as adapting and extending the model content. In his conceptualization of reuse oriented reference model design, vom Brocke identifies configuration, instantiation, specialization, aggregation, and analogy as particularly relevant [2]. As we base our work on this contribution, these are the principles we examine here. Others such as modification are specified in [7] and further elaborated in the discussion section.

- Configuration: Model parts are adopted according to the parametrization of the process domain. Individual model parts are selected and derived from a configurable component.
- Instantiation: General domain aspects are designed as a framework providing generic placeholders for plugging in model parts, considering the requirements of the application domain.

- Specialization: Entire contents of a generic model are adopted into a specific model, allowing individual modification and extension. The resulting model contains all contents of the individual model.
- Aggregation: Contents delivered by various part models are adopted into the new model, composed and extended as necessary. The resulting model is composed of the individual model parts.
- Analogy: Seemingly similar solutions are employed in a creative way to tackle new problems. The individual model is used as orientation for the design of the resulting model, such that they are perceived to be coinciding in certain aspects.

2.3 Reference Model Mining

Reference Model Mining describes the (semi-)automated derivation of a reference model from a set of individual models by identifying commonalities in a set of input models and constructing a new model on that basis (as illustrated in Figure 2).

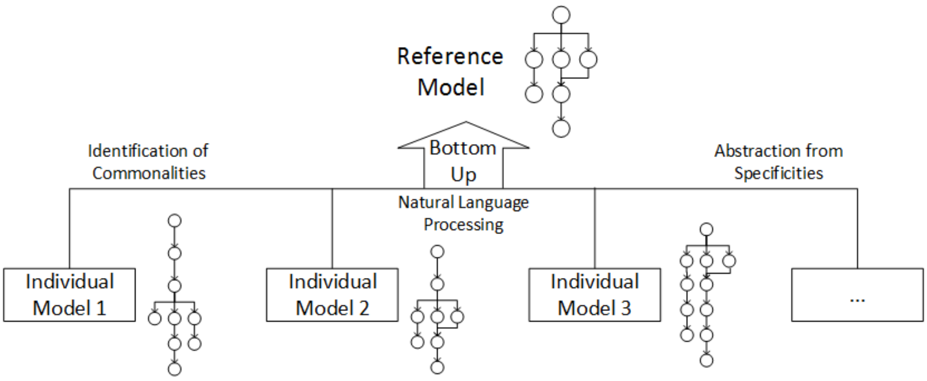


Figure 2: Reference Model Mining

3 Related Work

The concept of situational reference model construction based on design principles is not new, but has yet only been elaborated for deductive reference model development [1, 2]. Inductive reference model development has only recently gained attention in research, so there is little methodological seminal work. Fettke defines a seven-stage-framework for reference modeling methods [8]. First ideas towards S-RMM are presented in [9], where the choice of an appropriate mining technique is discussed.

A number of contributions describe concrete techniques and approaches to Reference Model Mining, but do not take on a methodological perspective, reflecting on the ways of model construction and the requirements of specific use cases. Process variants may either be mined in relation to an existing reference model or without one [10]. Different similarity measures, such as frequents common substructures [11] or heuristic approximations of the Graph-Edit Distance [12], are used to determine input model commonalities. Other approaches employ configurable process models [13], genetic algorithms [14, 15], or Process Model Abstraction [16]. While all of these contributions make the case for Reference Model Mining in general, none of them acknowledge the differences between existing approaches or indicate in which context their suggested technique would especially valuable. Different mining techniques employ different similarity measures (e.g. structural [12] or semantic [9]) and construction methods (e.g. deterministic [11] or heuristic [15]), resulting in differences between the mined reference models. In addition, due to restrictions on the input models, not every technique may be applied to every set of input models. Some contributions describe the influence of a certain parameter on the resulting reference model. For example, a frequency threshold as in [11, 12], will determine the model size and thus the character. The higher it is, the smaller and the more generic the resulting reference model. This is related to the underlying design principle, but not explicitly mentioned as such.

Some authors apply Situational Reference Model Mining by inductively developing reference models for a certain use case in a certain domain, without explicitly considering a generic procedure model or specific design principles [17-19]. Other techniques could generally be applied for Reference Model Mining, although that is not their primary use case. For example, Process Model Merging is primarily intended for process consolidation, but a consolidated model can also be interpreted as a reference model [20]. The same holds for Process Model Integration, especially in a hierarchical way [21]. If the reference model development is targeted towards certain quality aspects, it might make sense to choose it accordingly from process model configurations [22].

Our intention here is to extend the existing concept of RMM to consider the situational context, i.e. the intended target models, when choosing and executing a mining technique. Therefore, we want to create unified guidelines for S-RMM, which reference model designers can use for an easier and better application of Reference Model Mining. Depending on their individual use cases, designers should be able to make informed choices on their design principles and suitable mining techniques.

4 Situational Reference Model Mining

4.1 Idea and Conceptualization

Figure 3 describes the main idea behind S-RMM by extending and substantiating the reuse-oriented reference model design process from Figure 1. In Reference Model Mining, the reference model is automatically derived from a given set of input models, using a certain mining technique. The reference model is then used for the construction of the target models in a certain application context. Depending on the situational circumstances and the target model requirements, a certain design principle is applied to derive the target model from the reference model. This design principle poses certain restrictions and requirements to the reference model design, which is mainly influenced by the choice of the technique that was used to mine the reference model in the first place. Hence, the choice of mining technique is ultimately determined by the situational context of the reference model application.

Reference Model Mining itself is a way of constructing a reference model, i.e. an instantiation of Design For Reuse. Depending on the choice of mining technique, different techniques are used to determine the input model commonalities and construct the reference model. Hence, the choice of mining technique determines not only the content, but also the character of the reference model. It restricts the application of design principles for target model derivation, i.e. Design With Reuse, which is a use case for Reference Model Mining.

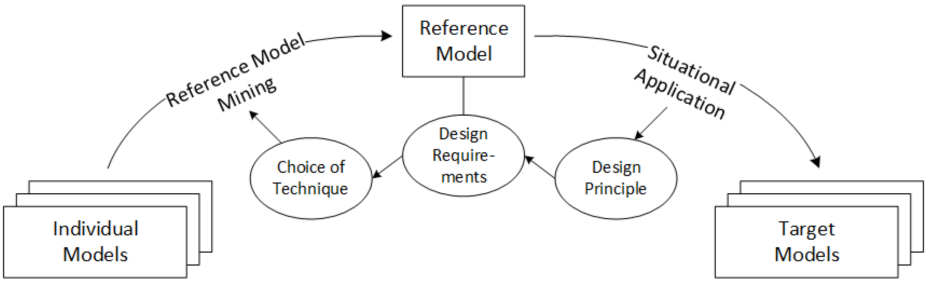


Figure 3: Main Idea behind Situational Reference Model Mining

4.2 Procedure Model

A general procedure model for Situational Reference Model Mining is shown in Figure 4. It is built around the conceptualization of S-RMM in Figure 3. The procedure model describes a generic execution process of an S-RMM application. It consists of ten steps, each of which belongs to one of the two generic design processes. DWR is concerned with the target model construction (i.e. the reference model application), whereas DFR focusses on the reference model construction (i.e. the actual mining). The generic S-RMM process starts with Design For Reuse, where seven steps are executed, and continues in Design With Reuse, with three steps.

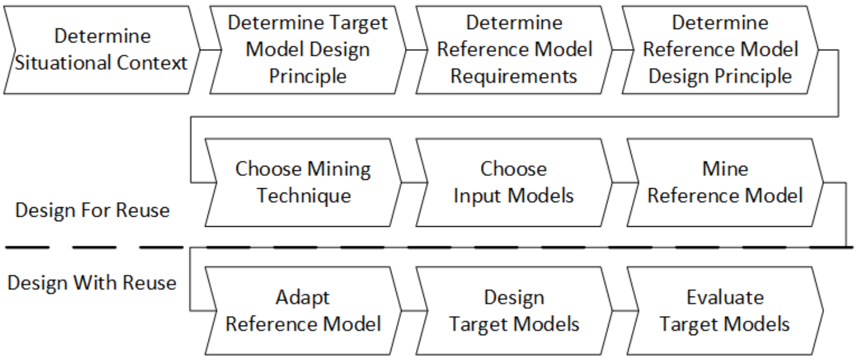


Figure 4: Procedure Model for Situational Reference Model Mining

1. Determine Situational Context: As a basis for any S-RMM application, the situational context has to be analyzed by determining the intended use for the target models and the required characteristics that follow from it. These may include defining the type of process, type of model, process domain, modeling language, level of abstraction, scope, or other target model characteristics. The situational context is also mainly influenced by the available input models, as their characteristics determine which design principles may be reasonably applied.
2. Determine Target Model Design Principle: Depending on the situational context and the inferred requirements, the target model design principle is chosen, based on the assessment which of the five principles satisfies the requirements the best.
3. Determine Reference Model Requirements: From the chosen design principle follow several requirements to the design of the reference model design. These are mainly independent from the situational context, as they follow mostly from the principle itself. For example, applying configuration requires a configurable reference model, while instantiation calls for generic process interfaces that can be individually specified.
4. Determine Reference Model Design Principle: Depending on the required design of the reference model, the process designer has to choose the design principle that should be applied in the mining process in order to fulfill these requirements.
5. Choose Mining Technique: Choosing an applicable and appropriate reference model mining technique is influenced by the chosen reference model design principle, but also the situational context that was previously analyzed, as the reference model has to fulfill a number of constraints. If several techniques qualify, it might be necessary to compare the resulting reference models to determine the best fit for the situational use case.
6. Choose Input Models: Usually, a set of input models is selected prior to beginning the mining process, as they determine the situational context. However, due to possible restrictions and requirements, the final set of input models can only be selected after the mining technique is chosen.
7. Mine Reference Model: The reference model is obtained by applying the mining technique to the chosen set of input models. As they depend on both the input data and the situational context, potential parameter configurations have to be individually determined to yield the best-fitting reference model.
8. Adapt Reference Model: As reference model mining techniques are usually fully automated, the resulting model may not fulfill all the requirements derived from the situational context. Hence, it may have to be manually adapted, for example by adding, deleting or renaming nodes, or complementing the reference model with deductively developed model parts.
9. Design Target Models: After the reference model is finalized, it can be used for the target model construction. Therefore, the design principle determined in step 2 is now applied to the reference model. Each target model undergoes a separate construction process, where the individual model requirements are addressed in the best possible way.
10. Evaluate Target Models: The goal of applying the S-RMM procedure model is to design a reusable and thus useful reference model and use this as a basis for high-

quality target models. Hence, in the last step, the target models are evaluated against the requirements derived from their intended use case. This step may lead to individual adaptations of the target models, but may also serve to enhance the reference model for further reusing. In addition, this step allows process designers to reflect on the S-RMM process as a whole, pointing out eventual improvements.

4.3 Analysis of Existing Mining Techniques

Table 1: Analysis of Existing Mining Techniques

Target Model Design Principle	Reference Model Requirements	Input Model Characteristics	Reference Model Design Principle	RMM Techniques
Configuration	Subsumes different characteristic values of a domain aspect (e.g. different process types).	Belong to the same domain, but describe different instantiations (e.g. subdomains).	Aggregation	[9-13, 20, 21]
			Analogy	[10, 14, 23]
Instantiation	Contains generic interfaces acting as placeholders for certain domain aspects (e.g. activities).	Come from the same domain, but differ in certain domain aspects that can be abstracted.	Aggregation	[16]
Specialization	Contains only universally applicable, contextually adaptable domain aspects.	Contain identical fragments representing universally valid domain aspects.	Aggregation	[9-12, 20]
			Analogy	[14, 23]
Aggregation	Requires several models, each covering one aspect relevant to the application context.	Originate from different backgrounds, but complement each other in terms of content.	--	
Analogy	Contains fragments that are directly applicable in target model context (i.e. in content and level of abstraction)	At least one is generic enough to serve as the basis for reference model design. The others can e.g. be derived variants.	Configuration	[13, 22]
			Aggregation	[9-12, 21]
			Analogy	[14, 23]

In order to provide a guideline for applying the procedure model for S-RMM, we analyze existing mining techniques regarding their underlying principles and requirements, as summarized in Table 1. For each target model design principle, we list the necessary reference model requirements and input model characteristics, such that this principle is applicable. Then, we suggest corresponding reference model design principles and, for each pair, a suitable mining technique. The analysis is restricted to those combinations of target and reference model design principle that are described in section 2.2 and actually substantiated by a mining technique, not those that generally exist or make sense.

Table 1 suggests both techniques that are specifically devised for Reference Model Mining (or Inductive Reference Model Development), such as Process Variant Clustering [10], and techniques that are originally intended for another use case, but can be employed accordingly, such as Process Model Merging [20]. Selection criteria were that the described technique (a) takes a set of models as input, (b) outputs a single model that is in some way based on the input models, (c) is fully automated, and (d) describes a domain-independent method that can be applied to any set of input models. This excludes methodical frameworks such as [8], partially manual approaches such as [17], or empirical, domain-specific reports such as [18]. The table is not intended as a complete list or a state-of-the-art analysis of Reference Model Mining, but as a complementary guideline to our procedure model.

5 Case Study

To illustrate the concept of Situational Reference Model Mining, we apply it in a case study. Figure 5 shows three auditing processes for three different types of retail business (wholesale, warehousing, central processing), translated and slightly adapted from their documentation in the Retail-H reference model [24]. These models will serve as both the target models and the individual models in our case study.

Determine Situational Context. Our target models treat the same process (auditing) in three different sub-domains of the retail domain, i.e. wholesaling, warehousing, and central processing. We assume that the objective is to design an inclusive reference model for auditing processes in retail, including all the differentiations in terms of business type. This means that the reference model should include the specificities of auditing in all three processes. Another option would be to design a model containing only those components that are present in all the input models, i.e. an excerpt.

Determine Target Model Design Principle. Since we intend to design an inclusive reference model, which represents several different sub-domains, the reference model scope is larger than the scope of the target models. Hence, parts of the reference model are irrelevant for each target model and should be deleted. Configuration appears to be a suitable design principle for the derivation of the reference model, as it allows for the selective adoption of applicable model parts and their adaption to the individual situation.

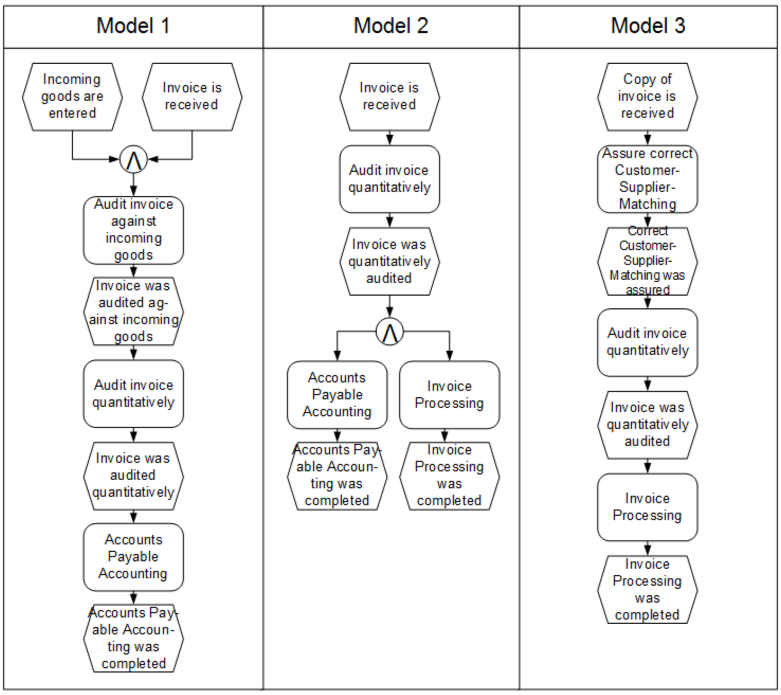


Figure 5: Example Models used for Case Study

Determine Reference Model Design Principle. Aggregation entails adopting and composing parts of several individual models, such that the resulting model subsumes the input models. Since all input models are equally considered, it is an applicable design principle for the reference model. If one input model were used as the basis for the reference model design, specialization or analogy would be more appropriate.

Choose Mining Technique. An inclusive reference model should be mined with a technique that preserves both the elements and the semantics of the individual model. The objective of Process Model Merging [20] is to construct a consolidated model out of a set of process models that share common fragments. As merged models are meant to subsume a set of process model variants, the mining technique is suited for our use case here.

Choose Input Models. For our case study, our input models are equivalent to the set of target models. As the merging algorithm is defined on Event-Driven Process Chains, it can be applied straightforwardly. A pairwise mapping between the input models is required as additional input. We assume that nodes carrying an identical label are mapped onto each other. Process Merging is defined on model pairs, however, the merging order should not have an influence on the resulting model.

Mine Reference Model. We mine the reference model by first merging model 1 and model 2 and then merging the resulting model with model 3. The resulting reference model is shown on the left of Figure 6.

Adapt Reference Model. Technically, the resulting reference model could be used for configuring the target models. However, the model is not syntactically correct, as an event ("Invoice is received") is followed by an XOR-connector. Hence, we adapt the model manually to include the differentiation between the warehousing and wholesale sub-domain. This way, we yield a syntactically correct EPC, which can be configured in the next step.
Design Target Models. The three target models are derived by configuring the adapted reference model. In our use case, this means that those model parts that are irrelevant for the individual sub-domain are removed from the model and the remaining relevant model parts are reordered and connected to form a valid EPC. If applicable, OR-connectors should be configured to be semantically precise.
Evaluate Target Models. In our use case, the target models were predefined, so there is no need for an evaluation in terms of process implementation. However, we can state that both chosen design principles were applicable to the use case. Configuration and aggregation are a suitable pair, but configuration could also be matched with specialization or analogy.

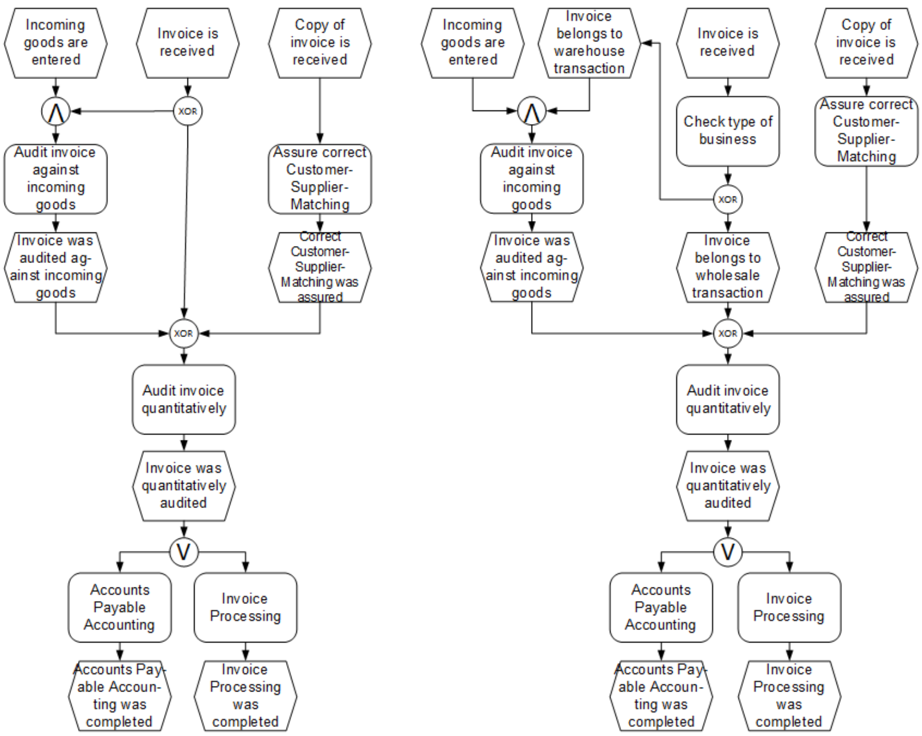


Figure 6: Merged reference model

6 Discussion and Conclusion

In this contribution, we introduce the concept of Situational Reference Model Mining. Based on the idea that the requirements to a reference model and thus the optimal mining technique depend on the situational context, we intend to give reference model designers a guideline for choosing the right mining technique for their individual purpose. As this idea has not yet been elaborated in the respective literature, our procedure model is intended as a first recommendation for a concrete approach. This way, we intend to increase the practical applicability of Reference Model Mining and make its advantages available to a wider range of users.

Our first assessment of S-RMM allows for several observations. First, different mining techniques yield different reference models, both in content and design, even when they are applied to the same set of input models. Parametrization may or may not be a decisive factor in reference model construction. The influence of parameters on the reference model contents and design has to be determined individually for each technique. For example, the order in process model merging should not influence the resulting model, while a frequency threshold (as for example in [11, 12]) determines whether a reference model is the intersection or the union of the input models. In this case, the parameter value also determines the design principle. Depending on the relative frequency of the reference model elements, configuration, analogy, or specialization are applicable. This is why multiple mining techniques apply to several combinations of target and reference model design principle, as seen in Table 1.

The results of the case study in section 5 underline our initial characterization of S-RMM. In order to apply Process Model Merging in a contextually meaningful way, we had to make a number of assumptions, such as the intended use of the reference model, the target models, and the related design principles. While all assumptions made above form a reasonable use case for reference modeling, there exists a plethora of other potential use cases, where the same input models would yield completely different results. Also, we saw that the combination of target and reference model design principle is not sufficient to choose an appropriate mining technique. As shown in Table 1, Process Model Merging is only one of many applicable techniques for the combination of aggregation/configuration. Applying another technique might not yield a reference model as the union of the input models, as seen in Figure 6. Nevertheless, it might be a meaningful reference model in a number of use cases.

Our analysis of existing mining techniques in Section 3.3 shows that the aggregation principle is predominant in reference model construction, while configuration is the mainly followed principle in reference model application. This is not surprising, given the nature of reference model mining. When deriving a reference model with a certain degree of universal applicability from a set of input models, aggregating their common features is evident but not the only technique to achieve a meaningful model. Basing the reference model on an input model and adapting it to reduce the difference to the remaining ones realizes the analogy principle, as for example in [10]. On the other hand, a reference model that aggregates aspects from different sub-domains has to be configured in order to obtain a model that applies to only one of them. However, aggregation may also yield the most common fragments (such as Process Model Intersection in [20]), where specialization or instantiation are appropriate design principles. To conclude, although aggregation/configuration is prominent, the pairwise combination of construction and application principle is not automatically given, but depends on the characteristics of the mining technique.

The analysis in Section 3.3 also shows that the instantiation principle is underrepresented in Reference Model Mining. Only if the reference model is constructed by means of process abstraction, the target models may be derived by means of instantiation. This is due to the fact that most existing approaches to Reference Model Mining are not capable of handling input models with varying degrees of abstraction. Hence, the abstraction level remains the same across all the input models and the reference model. The generic placeholder elements necessary for instantiation cannot be derived from differing, but more specific input models.

Our analysis also reveals that currently there exists no applicable technique for deriving the reference model by means of aggregation. That is because aggregation draws on several conceptual models covering different aspects of the situational context that are to be composed in the target model. None of the existing mining techniques is explicitly set out to mine several different reference models covering different aspects of the defined domain. However, such a scenario is realistic, for example when the reference model is supposed to cover a large domain, which should be divided into sub-domains to ensure the reference model applicability.

In this contribution we draw on the five principles configuration, instantiation, specialization, aggregation and analogy, as defined in [2]. However, these are not the only principles to be considered for reference model design. For example, Delfmann suggests modification as another design principle, allowing all changes to the reference model that do not result in erroneous or inconsistent models [7]. Besides that, principles like elimination or union might also be useful for reference model design. Elimination would allow designers to delete unnecessary elements from a reference model, whereas union would merge several models, without aggregating their contents.

Our analysis of existing mining techniques in Table 1 also acts as a gap analysis, identifying further research potentials and objectives and allowing for a more structural design of new mining techniques. The main motivation for this contribution is to increase the practical applicability of Reference Model Mining. Currently, there exist a number of publications that focus on technical and methodical aspects, as well as a few implementations, but few concrete suggestions for their application. By coining the term “Situational Reference Model Mining”, we emphasize that the choice of technique is relevant, i.e. they cannot always be interchangeably used. The procedure model, in combination with the analysis of existing techniques, is supposed to be a guideline for both reference modeling researchers and practitioners. However, it has not yet been evaluated by being applied in a large-scale context. Gaining more experience in practical applications of existing RMM techniques remains one of the major objectives of further reference modeling research. Our underlying assumptions, however, should be critically assessed. For example, in some cases it could make sense to develop situationally adequate target models instead of choosing an appropriate the mining technique.

Acknowledgement: The research described in this paper was partly supported by a grant from the German Research Foundation (DFG), project name: “Konzeptionelle, methodische und technische Grundlagen zur induktiven Erstellung von Referenzmodellen (Reference Model Mining)”, support code GZ LO 752/5-1. The authors would also like to thank the anonymous reviewers for their valuable comments which helped to improve this paper.

References

1. vom Brocke, J.: Referenzmodellierung – Gestaltung und Verteilung von Konstruktionsprozessen. Wirtschaftswissenschaftliche Fakultät. Universität Münster, Münster (2002)
2. vom Brocke, J.: Design Principles for Reference Modeling: Reusing Information Models by Means of Aggregation, Specialisation, Instantiation, and Analogy. In: Fettke, P., Loos, P. (eds.) Reference Modeling for Business Systems Analysis, pp. 47-75. Idea Group Publishing, Hershey (2007)
3. Fettke, P., Loos, P.: Perspectives on Reference Modeling. In: Fettke, P., Loos, P. (eds.) Reference Modeling for Business Systems Analysis, pp. 1-20. IGI Publishing, Hershey, PA (2007)
4. vom Brocke, J., Fettke, P.: Referenzmodellierung. In: Kurbel, K., Becker, J., Gronau, N., Sinz, E.J., Suhl, L. (eds.) Enzyklopädie der Wirtschaftsinformatik - Online-Lexikon, vol. 2015. Oldenbourg Wissenschaftsverlag, Munich, Germany (2013)
5. Rosemann, M., van der Aalst, W.M.P.: A Configurable Reference Modelling Language. Information Systems 32, 1-23 (2007)
6. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for Information Systems Research. Journal of Management Information Systems 24, 45-77 (2007)
7. Delfmann, P.: Adaptive Referenzmodellierung. Methodische Konzepte zur Konstruktion und Anwendung wiederverwendungsorientierter Informationsmodelle. Logos Verlag, Berlin (2006)
8. Fettke, P.: Eine Methode zur induktiven Entwicklung von Referenzmodellen. In: Kundisch, D., Suhl, L., Beckmann, L. (eds.) Multikonferenz Wirtschaftsinformatik 2014 (MKWI), pp. 1034-1047. GITO-Verlag, Paderborn (2014)
9. Rehse, J.-R., Fettke, P., Loos, P.: An Execution-Semantic Approach to Inductive Reference Model Development. 24th European Conference on Information Systems (ECIS). Association for Information Systems (AIS), Istanbul, Turkey (2016)
10. Li, C., Reichert, M., Wombacher, A.: Mining business process variants: Challenges, scenarios, algorithm. Data Knowl. Eng. 70, 409-434 (2011)
11. Rehse, J.-R., Fettke, P., Loos, P.: A graph-theoretic method for the inductive development of reference process models. Software & Systems Modeling. Springer, Berlin et al. (2015)
12. Ardalani, P., Houy, C., Fettke, P., Loos, P.: Towards a Minimal Cost of Change Approach for Inductive Reference Model Development. 21st European Conference on Information Systems (ECIS 2013), vol. Paper 127. AIS, Utrecht, Netherlands (2013)
13. Gottschalk, F., van der Aalst, W.M.P., Jansen-Vullers, M.H.: Mining Reference Process Models and Their Configurations. In: Meersman, R., Tari, Z., Herrero, P. (eds.) On the Move to Meaningful Internet Systems: OTM 2008 Workshops, OTM Confederated International Workshops and Posters, ADI, AWeSoMe, COMBEK, EI2N, IWSSA,

MONET, OnToContent + QSI, ORM, PerSys, RDDS, SEMELS, and SWWS 2008, Monterrey, Mexico, November 9-14, 2008, pp. 263-272. Springer, Berlin et al. (2008)
14. Martens, A., Fettke, P., Loos, P.: A Genetic Algorithm for the Inductive Derivation of Reference Models Using Minimal Graph-Edit Distance Applied to Real-World Business Process Data. In: Kundisch, D., Suhl, L., Beckmann, L. (eds.) Multikonferenz Wirtschaftsinformatik 2014 (MKWI), pp. 1613-1626. GITO-Verlag, Paderborn (2014)
15. Yahya, B.N., Bae, H., Bae, J., Kim, D.: Generating Business Process Reference Model using Genetic Algorithm. Biomedical Fuzzy Systems Association, vol. BMFSA 2010, pp. 245-248, Kitakyushu, Japan (2010)
16. Rehse, J.-R., Fettke, P., Loos, P.: Eine Untersuchung der Potentiale automatisierter Abstraktionsansätze für Geschäftsprozessmodelle im Hinblick auf die induktive Entwicklung von Referenzprozessmodellen. In: Alt, R., Franczyk, B. (eds.) 11th International Conference on Wirtschaftsinformatik (WI2013), vol. 2, pp. 1277-1291, Leipzig (2013)
17. Gröger, S., Schumann, M.: Entwicklung eines Referenzmodells für die Gestaltung des Drittmittel-Prozesses einer Hochschule und Ableitung von Einsatzgebeiten für Dokumenten- und Workflow-Management-Systeme. Georg-August-Univeristät (2014)
18. Karow, M., Pfeiffer, D., Räckers, M.: Empirical-Based Construction of Reference Models in Public Administrations In: Bichler, M., Hess, T., Krcmar, H., Lechner, U., Matthes, F., Picot, A., Speitkamp, B., Wolf, P. (eds.) Proceedings of the Multikonferenz Wirtschaftsinformatik 2008. Referenzmodellierung, pp. 1613-1624. GITO-Verlag, München (2008)
19. Aier, S., Fichter, M., Fischer, C.: Referenzprozesse empirisch bestimmen − Von Vielfalt zu Standards. Wirtschaftsinformatik & Management 3, 14-22 (2011)
20. La Rosa, M., Dumas, M., Uba, R., Dijkman, R.M.: Business Process Model Merging: An Approach to Business Process Consolidation. ACM Transactions on Software Engineering and Methodology (TOSEM) 22, (2013)
21. Fettke, P.: Integration von Prozessmodellen im Großen: Konzept, Methode und experimentelle Anwendungen. In: Thomas, O., Teuteberg, F. (eds.) 12th Interntational Conference on Wirtschaftsinformatik, pp. 453-467, Osnabrück, Germany (2015)
22. Schunselaar, D.M., Verbeek, E., Reijers, H.A., van der Aalst, W.M.: Using Monotonicity to Find Optimal Process Configurations Faster. In: 4th International Symposium on Datadriven Process Discovery and Analysis, pp. 123-137. Citeseer, (Year)
23. Yahya, B.N., Bae, H., Bae, J., Kim, D.: Generating valid reference business process model using genetic algorithm. International Journal of Innovative Computing, Information and Control 8, 1463-1477 (2012)
24. Becker, J., Delfmann, P., Knackstedt, R., Kuropka, D.: Konfigurative Referenzmodellierung. In: Becker, J., Knackstedt, R. (eds.) Wissensmanagement mit Referenzmodellen. Konzepte für die Anwendungssystem- und Organisationsgestaltung, pp. 25-144. Springer, Berlin et al. (2002)

Collaborative Business Process Management – A Literature-based Analysis of Methods for Supporting Model Understandability

Andreas Hermann, Hendrik Scholta, Sebastian Bräuer, Jörg Becker

University of Münster, ERCIS, Münster, Germany
{firstname.lastname}@ercis.uni-muenster.de

Abstract. Due to the growing amount of cooperative business scenarios, collaborative Business Process Management (cBPM) has emerged. The increased number of stakeholders with minor expertise in process modeling leads to a high relevance of model understandability in cBPM contexts. Despite extensive works in the research fields of cBPM and model understandability in BPM, there is no analysis and comprehensive overview of methods supporting process model understandability in cBPM scenarios. To address this research gap, this paper presents the results of a literature review. The paper identifies concepts for supporting model understandability in BPM, provides an overview of methods implementing these concepts, and discusses the methods' applicability in cBPM. The four concepts process model transformation, process model visualization, process model description, and modeling support are introduced. Subsequently, 69 methods are classified and discussed in the context of cBPM. Results contribute to revealing existing academic voids and can guide practitioners in cBPM scenarios.

Keywords: Business Process Management, Collaborative Business Process Management, Model Understandability, Literature Review

1 Introduction

Business Process Management (BPM) is a discipline that combines computer science and management science and has gained a considerable amount of attention over the last decades [1, 2]. The growing importance of cooperation due to globalization and the trend of blurring organizational boundaries lead to collaboration in BPM [1, 3–5]. *Collaborative Business Process Management* (cBPM) is concerned with the management of business processes across organizational boundaries [4]. Since cBPM integrates different collaborating organizations [5, 6], the number of stakeholders involved in business process modeling activities is high. However, since not all relevant stakeholders are experienced in process modeling and the particular notations [7], the models might not be fully understood by all stakeholders [8]. Clearly, there exists a gap between modeling experts and inexperienced stakeholders like domain experts [8–10]. This gap needs to be bridged to guarantee success in cBPM projects [11–13].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Hermann, A.; Scholta, H.; Bräuer, S.; Becker, J. (2017): Collaborative Business Process Management - A Literature-based Analysis of Methods for Supporting Model Understandability, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 286-300

For addressing this knowledge gap, a considerable amount of research in the field of business process model understandability has been conducted [14, 15]. Many of these contributions are experimental works focusing on factors that influence model understandability [14]. Researchers investigate factors of modeling languages, model characteristics, model content related factors or personal factors that influence model understandability [7, 16–19]. For example, Figl et al. investigate the influences of routing symbols of modeling languages on process model understandability [19]. Reijers and Mendling observed that model characteristics such as the average connectors degree or the overall density of a model affect its understandability [7]. Furthermore, Mendling and Strembeck found out that long element labels influence model understandability negatively [16].

Researchers recognize the relevance of cBPM on the one hand and model understandability in the context of BPM on the other hand. However, there is currently no review that analyzes methods implementing concepts for supporting model understandability in BPM and evaluates the methods' suitability for cBPM. This paper's objective is to identify existing methods in the intersection of the topics model understandability and BPM and to discuss the general applicability of these methods in the context of cBPM. Therefore, we have performed a structured literature analysis and address the following research question: *What are methods supporting model understandability in BPM and to what extent are they applicable to cBPM?* Our results provide a comprehensive overview of the last decade's academic work on the topic of model understandability in BPM and a discussion on the usage in cBPM. Practitioners can use it as guidance for identifying potential methods supporting their business. Academics can rely on our work to identify academic voids and plan future research.

The remainder of this paper is structured as follows: Section 2 presents the underlying research background on cBPM and model understandability. In Section 3, we describe the applied research method. We derive concepts for supporting model understandability and provide an overview of methods that implement these concepts in Section 4. In Section 5, we combine model understandability and cBPM by discussing the methods' applicability for cBPM. Finally, in Section 6, we draw a conclusion and present potential future work.

2 Research Background

2.1 BPM and cBPM

BPM is concerned with operational business processes and their management, improvement [2], re-design, analysis, or support with information systems (IS) [18]. An increasing importance of global value chains leads to a trend of blurring organizational boundaries in the context of BPM [3, 4]. Against this background, cBPM is an expansion of BPM that strives to cover business processes across inter-organizational boundaries [20].

cBPM can be described as the management of (collaborative) business processes across organizational boundaries involving actors from inside or from outside an organization [4]. Hence, in comparison to traditional BPM, cBPM incorporates an *increased number of stakeholders* since it affects not only a single organization but also at least one additional organization [5]. Besides, since organizations with a similar business model rather tend to compete than to collaborate, the *group of stakeholders possesses a high degree of heterogeneity.*

Consequently, the collaboration comprises organizations with different product portfolios and from different domains, which strive to deliver more value to their customers. Due to the increased number of stakeholders, the higher degree of heterogeneity, and the resulting need for coordination, cBPM has to cope with *more complex business processes* [3]. Increased complexity in the execution of business models also affects the modeling and thus results in *more complex process models* in cBPM [3]. Process models have to capture the more sophisticated control flow relations in these business processes and have to integrate different modeling conventions in the participating organizations [3].

Additionally, in contrast to traditional BPM, *privacy* plays a more important role in cBPM since confidential information of one organization must not cross organizational boundaries [20]. Therefore, certain information in process models must be kept secret to cooperating organizations [20].

In consequence, it is especially challenging in the context of cBPM to reach a common understanding of the process models among relevant stakeholders [21]. Although model understandability is relevant in traditional BPM, the mentioned reasons increase its importance in cBPM, but likewise, impede its achievement.

2.2 From Model Quality to Model Understandability

Business process models as central artifacts in BPM [2, 5] are the basis for the development of process-oriented IS [22, 23]. Process models have to possess high quality to obtain IS of high quality [24–26]. Model understandability can be considered as a factor of model quality [14, 15]. Some quality frameworks [27, 28] include the dimension *pragmatic quality*. Pragmatic quality is concerned with the degree to which a model is correctly interpreted or understood by an end-user or stakeholder [27, 29]. Accordingly, understandability is often referred to as a factor of pragmatic quality [14, 15, 25, 30, 31].

Pragmatic quality is defined as the "correspondence between the model and the […] interpretation of the model" [29, p. 94]. So-called *pragmatic means*, introduced by Lindland et al. [27], can be applied to reach the goal of pragmatic quality, i.e. understanding a model. In this sense, pragmatic means make a model more understandable [27, 29]. Pragmatic means are model animation, model simulation, model visualization, model transformation, model filtering, model abstraction, model translation, model explanation, as well as aesthetics for diagram layout, model paraphrasing, and participant training [27, 32]. Based on pragmatic means, we derive concepts that support model understandability and analyze methods that implement these concepts.

3 Research Method

The literature-based analysis conducted in this work is based on the approaches for systematic literature reviews proposed by Webster and Watson [33] and vom Brocke et al. [34]. To further define the scope [34] and to articulate the contribution of the work in detail [33], vom Brocke et al. [34] propose the application of the taxonomy for literature reviews by Cooper [35]. The paper's taxonomy is visualized in Table 1. The gray cells in the table below represent the focus of this literature analysis.

Table 1. Classification (gray cells) of the present literature-based analysis following [34]

<table>
<tr><th>Characteristic</th><th colspan="6">Categories</th></tr>
<tr><td>focus (1)</td><td>research outcomes</td><td colspan="2">research methods</td><td colspan="2">theories</td><td>applications/practices</td></tr>
<tr><td>goal (2)</td><td colspan="2">integration</td><td colspan="2">criticism</td><td colspan="2">central issues</td></tr>
<tr><td>organization (3)</td><td colspan="2">historical</td><td colspan="2">conceptual</td><td colspan="2">Methodological</td></tr>
<tr><td>perspective (4)</td><td colspan="3">neutral representation</td><td colspan="3">espousal of position</td></tr>
<tr><td>audience (5)</td><td>specialized scholars</td><td colspan="2">general scholar</td><td colspan="2">practitioners</td><td>general public</td></tr>
<tr><td>coverage (6)</td><td>exhaustive</td><td colspan="2">exhaustive and selective</td><td colspan="2">representative</td><td>central/pivotal</td></tr>
</table>

This paper's search process follows the guidelines for literature reviews as proposed by vom Brocke et al. [34]. The approach includes the four phases: journal search, database search, keyword search and backward and forward search. Webster and Watson [33] propose a topic-based search across all relevant journals. Since it includes a large number of electronic articles and provide access to leading IS journals, Elsevier Scopus was selected as database.

To search for relevant publications, the search string in Figure 1 was used[1]. The search string comprises three constituents: Terms from the cBPM literature, terms related to pragmatic means respectively model understandability, and additional terms that are used to further limit the scope on process modeling respectively conceptual modeling. The search string was applied on the 24th October of 2016 and led to 2448 results.

After this keyword search, the results were evaluated regarding their relevance [34]. For this purpose, vom Brocke et al. [34] propose an analysis of the titles, abstracts or full texts of the search results. A title-based analysis of the total results led to 102 results considered as relevant. Based on their abstracts, these 102 publications were then analyzed in detail concerning their relevance for answering the research question. This procedure led to 43 relevant methods from 43 publications.

In addition to a keyword search, Webster and Watson [33] recommend a forward and backward search based on the evaluated results of the keyword search. Using the results of the keyword search, a one-level backward and forward search was conducted, which included referencing and referenced works of the 43 publications that were

---

[1] The search string uses the syntax of Elsevier Scopus. It includes the Boolean operators OR and AND. The * is a wildcard symbol. The search was limited to the following subject areas which are considered as being relevant: Computer Science, Engineering, Mathematics, Decision Sciences, Multidisciplinary and Business, Management and Accounting.

considered as relevant. Performing the backward search and the forward search with the database Google Scholar led to 26 additional methods from 14 publications. Our literature search finally resulted in 69 methods from 57 publications that were considered relevant.

(TITLE-ABS-KEY("collaborat" OR "choreograph*" OR "modeling support" OR "modelling support" OR "cooperat*" OR "interorganizational" OR "inter-organizational" OR "cross-organizational" OR "filter" OR "view" OR "filtering" OR "visual*" OR "translat*" OR "transform*" OR "layout" OR "training" OR "workshop" OR "explanation" OR "paraphras*" OR "simulat*" OR "execution" OR "animation") AND TITLE-ABS-KEY("pragmatic quality" OR "clarity" OR "interpretation" OR "understand*" OR "comprehen*") AND TITLE-ABS-KEY("process model*" OR "conceptual model*")) AND SUBJAREA(MULT OR COMP OR ENGI OR MATH OR BUSI OR DECI) AND (EXCLUDE(SUBJAREA,"ENVI") OR EXCLUDE(SUBJAREA,"EART") OR EXCLUDE(SUBJAREA,"SOCI") OR EXCLUDE(SUBJAREA,"MEDI") OR EXCLUDE(SUBJAREA,"PSYC") OR EXCLUDE(SUBJAREA,"MATE") OR EXCLUDE(SUBJAREA,"CENG") OR EXCLUDE(SUBJAREA,"BIOC") OR EXCLUDE(SUBJAREA,"AGRI") OR EXCLUDE(SUBJAREA,"PHYS") OR EXCLUDE(SUBJAREA,"CHEM") OR EXCLUDE(SUBJAREA,"ENER") OR EXCLUDE(SUBJAREA,"NEUR") OR EXCLUDE(SUBJAREA,"ARTS") OR EXCLUDE(SUBJAREA,"ECON") OR EXCLUDE(SUBJAREA,"HEAL") OR EXCLUDE(SUBJAREA,"IMMU") OR EXCLUDE(SUBJAREA,"NURS") OR EXCLUDE(SUBJAREA,"PHAR"))*

Figure 1. Scopus search string

4 Supporting Model Understandability in BPM

4.1 Concepts for Supporting Model Understandability in BPM

Following the search process, the relevant results were synthesized and analyzed using a concept-centric matrix [33]. The concepts for this matrix were derived in two different ways: Deductively with pragmatic means and inductively based on the search results themselves. Since pragmatic means are instruments making a model more understandable, they represent appropriate concepts for classifying the search results. For clarity and for facilitating a better discrimination between the concepts, similar pragmatic means were grouped together. The relevant concepts for this literature analysis that were derived from pragmatic means in this manner are *process model transformation*, *process model visualization*, and *process model description.* Besides, the search results led to some results that are best classified as the concept *modeling support*. The pragmatic mean *participant training* stretches across all concepts. Figure 2 visualizes the concepts.

Modeling support (I). The concept *modeling support* is not directly derived from any pragmatic means. However, the analysis of the search results led to a number of methods that are best assigned to this additional concept. Correspondingly classified methods strive to support model understandability already during the construction process of the model. Consequently, this concept comprises contributions providing a new or extended modeling language [36], a special modeling tool [37], or a method that uses existing modeling notations in an innovative way [38]. Methods that integrate the use of collaborative technologies such as commenting, audio-communication, video-communication, and chatting functionalities into the modeling process are also assigned to this concept [11]. In contrast to this, methods that operate on already constructed models are not part of this concept.

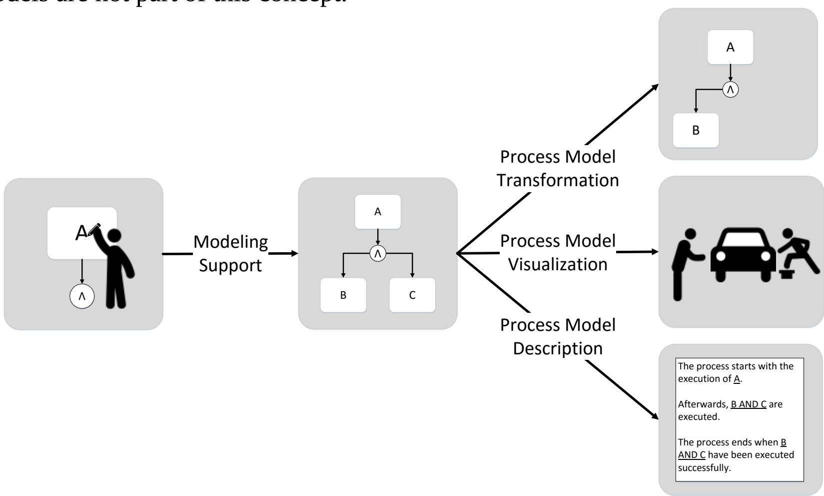


Figure 2. Concepts for supporting model understandability

Process model transformation (II). The concept *process model transformation* is derived from the pragmatic means *model transformation* and *aesthetics for diagram layout.* The means *model abstraction* and *model filtering* are also the basis for this concept since they are concerned with abstracting a model [29] or filtering out irrelevant model elements. This concept aims at the generation of specific views on models. As a result, transformation methods reduce model complexity [39], support the overall model understandability, and hence facilitate activities like the communication of the model to involved stakeholders [40].

Process model visualization (III). The pragmatic means *visualization*, *simulation*, and *animation* are the basis for the concept *process model visualization*. This concept comprises the alternative visualization of a process model's content. In other words, elements of a process model are depicted by alternative visual representations of non-model elements that substitute the original model elements. This concept also covers process animations. The use of visualization mechanisms enables an improved and understandable process model representation [9, 10, 41]. Whereas model transformation (II) relies on restructuring or visualization with alternative model elements, model visualization makes use of non-model elements.

Process model description (IV). *Process model description* comprises the pragmatic means *model paraphrasing*, *explanation*, and *translation* [27, 32]. This concept covers textual descriptions or explanations of a process model to raise its understandability. This, for instance, includes the automatic generation of textual descriptions capturing the process logic as depicted in a process model. The generation of natural language process model descriptions with explanatory character supports model understandability as it allows focusing on process model semantics rather than syntax. In contrast to the concepts (II) and (III), this concept makes use of texts and does not incorporate graphical elements.

4.2 Methods for Supporting Model Understandability in BPM

The identified literature was classified using the concepts explained above. The classification is presented in a concept-centric matrix (see Table 2) as proposed in [33]. The previously introduced concepts are not disjunctive. Consequently, a method can be assigned to more than one concept.

Modeling support. In total, there are 35 methods supporting the modeling process, which can be divided into three different groups, namely a) tools using collaborative technologies (No. 31-33, 38-53), b) methods providing new or extended languages (No. 9, 11-13, 16, 65) and c) tools using existing modeling languages (No. 6, 8, 17, 19, 29, 55-56, 58, 60, 63). Tools using collaborative technologies facilitate process modeling with the support of collaborative technologies such as commenting or text- and/or audio-based chats for supporting model understandability. Some methods of this concept focus on new modeling languages claiming to be less complex and easy to understand. Methods using existing modeling languages try to support model understandability by employing those languages in specific ways or adapting some existing modeling methods for process modeling.

Process model transformation. In total, the literature search led to 41 methods that are concerned with process model transformation. 22 out of 41 methods focus on altering the model's physical structure (No. 1-7, 10, 11, 14, 15, 18, 23, 25, 30, 42, 49, 51-53, 64). Generally speaking, these methods abstract from insignificant process model information, i.e. they focus on relevant information and omit irrelevant parts of a process model. In contrast, 28 out of 41 methods change the model's presentation, i.e. its appearance, scheme or layout (No. 6, 14, 19, 25-27, 31-36, 42, 48, 51-53, 56-62, 66-69). A transformation of a model's presentation does not change its physical structure to highlight relevant model information [39].

Process model visualization. The concept-centric matrix indicates that nine methods enable some kind of process model visualization. In this set of methods, the simplest form of process model visualization is accomplished by the use of additional non-model images (No. 20, 26). Advanced methods (No. 33, 31, 28, 32, 55) make use of 3D virtual world environments for visualizing a process model in a real-world like representation. The visualization of the token flow (No. 37, 66) in a process model is also a relevant implementation of this concept. The displayed token flow represents the execution order of activities to aid the analysis or validation of a process model.

Process model description. Five out of 69 methods are concerned with the generation and integration of process model descriptions capturing the process logic. Methods of this concept differ in the modeling language support and in the fashion, how they create process model descriptions. Methods can be generic as well as modeling language specific. Starting with a process model, descriptions are generated and integrated either automatically based on sophisticated algorithms (No. 21-22, 25, 54) or manually following specific guidelines (No. 24).

Table 2. Concept matrix

#	Src	Method	I	II	III	IV
1	[61]	Shared Process Model		•		
2	[62]	Tailored process model abstraction		•		
3	[63]	Process model abstraction slider		•		
4	[64]	Process model abstraction with data support		•		
5	[65]	Graph-based process model refactoring (IBUPROFEN)		•		
6	[66]	Modeling with Concurrent Task Trees	•	•		
7	[67]	Semantic process model abstraction		•		
8	[68]	Ontological process modeling	•			
9	[69]	Subject-oriented business process management	•			
10	[70]	Aspect-oriented BP Modeling (AO4BPMN)		•		
11	[36]	Knowledge intensive process notation	•	•		
12	[71]	Domain specific modeling languages (DSLs4BPM)	•			
13	[72]	Lightweight process modeling	•			
14	[73]	Deontic BPMN		•		
15	[74]	Scenario-based process modeling (GRETA)		•		
16	[75]	Role-based process modeling	•			
17	[76]	Context-driven process modeling	•			
18	[77]	Aspect-oriented BPM (AO-BPM)		•		
19	[78]	Spreadsheet-based process modeling	•	•		
20	[60]	end-user approach to process modeling			•	
21	[79]	Hypertext Expl. for process models (OMT Explainer)				•
22	[8]	Text Generation for process models				•
23	[80]	Personalized process model visualization (Proviado)		•		
24	[81]	Integration of natural language artifacts				•
25	[82]	Personalized and verbalized PM descriptions (proView)		•		•
26	[40]	Storyboard Augmentation		•	•	
27	[83]	3D repr. of business process models (RELAX NG)		•		
28	[9]	3D representation of BP Models			•	
29	[38]	Tangible business process modeling	•			
30	[84]	Event-driven Proc. Chain abstraction		•		
31	[85]	cBPM in 3D virtual worlds (I)	•	•	•	
32	[11]	cBPM in 3D virtual worlds (II)	•	•	•	
33	[10]	Modeling in 3D virtual worlds	•	•	•	
34	[86]	3D-Navigator for BP Models		•		
35	[87]	BPMN layout patterns		•		
36	[88]	Interactive BPMN layout tool		•		
37	[41]	Token flow animation			•	
38	cited in [12]	cPM (ProcessWave)	•			
39	[37]	cPM (CoMoMod)	•			
40	[42]	cPM (COMA)	•			
41	cited in [43]	cPM (BONAPART Collaborative)	•			
42	cited in [43]	cPM (Signavio Process Editor)	•	•		
43	cited in [43]	cPM (Enterrise Architect)	•			
44	cited in [43]	cPM (CA ERwin Process Modeler)	•			
45	cited in [43]	cPM (iGrafx Process Modeler)	•			
46	cited in [43]	cPM (Business Modeler Advanced)	•			
47	cited in [43]	cPM (Microsoft Visio)	•			
48	cited in [43]	cPM (Adonis)	•	•		
49	cited in [43]	cPM (Savvion Process Manager)	•			
50	cited in [43]	cPM (Innovator for Business Analysts)	•			
51	cited in [11]	cPM (ARIS Business Architect)	•	•		
52	cited in [11]	cPM (SAP StreamWork)	•	•		
53	cited in [11]	cPM (IBM Blueworks Live)	•	•		
54	[44]	Language generation for process models				•
55	[45]	Virtual business role-play	•		•	
56	[46]	Layout features of process models	•	•		
57	[47]	BPMN refactoring		•		
58	[48]	Layout features of process models	•	•		
59	[49]	Optimizing spaghetti process models		•		
60	[50]	Business processes to touch	•	•		
61	[51]	User-friendly model Visualizations		•		
62	[52]	Task-specific visual cues		•		
63	[53]	Collaborative articulation of models	•			
64	[54]	Collaborative product and process model		•		
65	[55]	Visualizing declarative process models	•			
66	[56]	Dynamic visualization of process models		•	•	
67	[57]	Revealing hidden dependencies		•		
68	[58]	Process model refactoring		•		
69	[59]	Syntax highlighting in process models		•		
Σ			**35**	**41**	**9**	**5**

5 Discussion

According to the research question and based on the characteristics of cBPM presented in the research background, we discuss the concepts and their methods for supporting model understandability regarding their applicability in cBPM.

Modeling support and cBPM. Due to the increased number of stakeholders in cBPM [4], collaborating techniques for modeling are required. Therefore, modeling methods that integrate collaborative technology are promising (e.g. No. 32, 39). However, these methods support the modeling process that is mainly performed by modeling experts instead of domain experts. Thus, the increased semantic complexity perceived by domain experts is not overcome and can only be addressed by communication between domain experts and modeling experts.

The group of methods that provide new or extended modeling language strives to increase model understandability by focusing on reduced syntactic complexity of the modeling language (e.g. No. 13). However, new modeling languages require extensive implementation efforts as existing modeling practices need to be reorganized which is especially relevant in cBPM settings with many diverse stakeholders.

Another group of methods focuses on existing modeling languages that are used in specific ways (e.g. No. 6, 29). Such methods could lead to a closer integration of modeling experts and domain experts during the model creation. An advantage is less implementation effort since these languages may already be in use.

Process model transformation and cBPM. Methods that modify the model's physical structure to abstract from insignificant details reduce the complexity of process models. In this way, the diverse domain knowledge of the many participating stakeholders [5] is addressed as personalized model views contain only relevant process logic. Furthermore, these methods are relevant to preserve autonomy and privacy of collaborating organizations, which are of increased importance in cBPM [89]. They allow the omission of sensitive and confidential internal information (e.g. No. 3, 23).

Transforming the process model's presentation allows for tailoring the model elements to the specifics of stakeholders. However, the increased model complexity in cBPM [3] is not completely addressed as the number of activities and control flow relations is not reduced by changing its presentation. Especially, the aspect of privacy issues in cBPM settings is not targeted by those methods since they do not hide model elements. Therefore, these methods are only applicable in combination with structural transformations to ensure privacy (e.g. No. 6, 14, 25).

Process model visualization and cBPM. The visualization of process models overcomes difficulties in understanding modeling language elements, i.e. the syntax of language elements [10]. Consequently, visualizations allow focusing on model semantics rather than syntax. To handle the increased semantic complexity in cBPM [3], visualization methods can be used since they abstract from model syntax and thereby decrease complexity for domain experts with low modeling expertise. In contrast, process model transformation methods rely on the model syntax and are therefore less appropriate in cBPM.

Since processes are more complex in cBPM scenarios [3], more complex methods are required for increasing model understandability in general. Hence, within the group of methods for process model visualization, 3D virtual environments are superior to less comprehensive methods (e.g. No. 31, 33). However, confidentially of private activities remains an open issue in all methods that are subsumed under this concept.

Process model description and cBPM. The set of methods that create process model descriptions strives to increase model understandability by natural language representations of process models. The translation of process models to natural language reduces the relevance of understanding formal modeling syntax. In consequence, these methods are more effective than methods for process model transformation to increase model understandability in cBPM in general.

Apparently, generic methods are superior to methods that are dedicated to a specific modeling language. Many diverse stakeholders with different modeling conventions are involved in a cBPM scenario. Therefore, methods that can handle different modeling languages are recommended (e.g. No. 22, 25). Besides, privacy requirements are more easily met with manual instead of automatic methods.

6 Conclusion

The aim of this paper was to derive concepts for supporting model understandability in the context of BPM, present exemplary implementations of these concepts, and discuss their suitability in cBPM contexts. The identified concepts *process model transformation*, *process model visualization*, *process model description* and *modeling support* are promising to be useful for cBPM-specific issues, although to varying degrees. Process model transformation is useful for specifying views to hide confidential information, process model visualization and process model description are suitable to increase the model understandability for domain experts with low modeling experience and methods supporting the modeling itself provide valuable mechanisms for the collaborative development of business process models. In total, 41 implementations for model transformation were detected; nine methods deal with model visualization, five methods focus on model description and 35 implementations provide modeling support.

This paper contributes to research by providing an overview of methods for addressing model understandability that allows for the identification of academic voids and presenting four categories to classify such methods. Practitioners can use our results as guidance for the use in cBPM scenarios. In future research, our paper can be extended by a detailed comparison of the analyzed methods related to each concept. Additionally, the methods can be empirically validated regarding their applicability in cBPM contexts. Besides, it can be investigated whether factors that influence model understandability differ from traditional BPM to cBPM.

Acknowledgements

This article was written in the BMBF-funded research project “Cooperation Experience” (funding label: 01XZ13013). We thank the project management agency DLR for their support.

References

1. Houy, C., Fettke, P., Loos, P., van der Aalst, W. M. P., Krogstie, J.: BPM-in-the-Large – Towards a Higher Level of Abstraction in Business Process Management. In: Janssen, M., Lamersdorf, W., Pries-Heje, J., and Rosemann, M. (eds.) E-Government, E-Services and Global Processes. pp. 233–244. Springer Berlin Heidelberg (2010).
2. van der Aalst, W. M. P.: Business Process Management : A Comprehensive Survey. ISRN Softw. Eng. 2013, 1–37 (2013).
3. Adam, O., Hofer, A., Zang, S.: A collaboration framework for cross-enterprise business process management. In: Preproceedings of the First International Conference on Interoperability of Enterprise Software and Applications INTEROP-ESA. pp. 499–510 (2005).
4. Niehaves, B., Plattfaut, R.: Collaborative business process management: status quo and quo vadis. Bus. Process Manag. J. 17, 384–402 (2011).
5. Aleem, S., Lazarova-Molnar, S., Mohamed, N.: Collaborative Business Process Modeling Approaches: A Review. In: Proc. of the 2012 IEEE 21st International workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises. pp. 274–279 (2012).
6. Chen, Q., Hsu, M.: Inter-enterprise collaborative business process management. In: Proceedings of the 17th International Conference on Data Engineering. pp. 253–260. Published by the IEEE Computer Society (2001).
7. Reijers, H. A., Mendling, J.: A Study Into the Factors That Influence the Understandability of Business Process Models. IEEE Trans. Syst. Man Cybern. Part A. 41, 449–462 (2011).
8. Leopold, H., Mendling, J., Polyvyanyy, A.: Generating Natural Language Texts from Business Process Models. In: Ralyte, J., Franch, X., Brinkkemper, S., and Wrycza, S. (eds.) Advanced Information Systems Engineering. pp. 64–79. Springer Berlin Heidelberg (2012).
9. Guo, H., Brown, R., Rasmussen, R.: Virtual worlds as a model-view approach to the communication of business processes models. CEUR Workshop Proc. 855, 66–73 (2012).
10. Brown, R. A.: Conceptual Modelling in 3D Virtual Worlds for Process Communication. In: Procedings of the 7th Asia-Pacific Conference on Conceptual Modelling (APCCM 2010). pp. 25 – 32. , Brisbane (2010).
11. Poppe, E., Brown, R., Recker, J., Johnson, D.: Improving Remote Collaborative Process Modelling using Embodiment in 3D Virtual Environments. In: 9th Asia-Pacific Conference on Conceptual Modelling. pp. 51–60 (2013).
12. Recker, J., Mendling, J., Hahn, C.: How collaborative technology supports cognitive processes in collaborative process modeling: A capabilities-gains-outcome model. Inf. Syst. 38, 1031–1045 (2013).
13. Recker, J., Reijers, H. A., van de Wouw, S. G.: Process model comprehension: The effects of cognitive abilities, learning style, and strategy. Commun. Assoc. Inf. Syst. 34, 199–222 (2014).
14. Houy, C., Fettke, P., Loos, P.: Understanding understandability of conceptual models - What are we actually talking about? In: Atzeni, P., Cheung, D., and Ram, S. (eds.) Conceptual Modeling. pp. 64–77. Springer Berlin Heidelberg (2012).
15. Houy, C., Fettke, P., Loos, P.: on the Theoretical Foundations of Research Into the Understandability of Business Process Models. In: Proceedings of the 22nd European Conference on Information Systems. pp. 1–26 (2014).
16. Mendling, J., Strembeck, M.: Influence Factors of Understanding Business Process Models. In: Abramowicz, W. and Fensel, D. (eds.) Business Information Systems. pp. 142–153. Springer Berlin Heidelberg (2008).
17. Mendling, J., Reijers, H. A., Cardoso, J.: What Makes Process Models Understandable? In: Alonso, G., Dadam, P., and Rosemann, M. (eds.) Business Process Management. pp. 48–63. Springer Berlin Heidelberg (2007).

18. Mendling, J., Strembeck, M., Recker, J.: Factors of process model comprehension-Findings from a series of experiments. Decis. Support Syst. 53, 195–206 (2012).
19. Figl, K., Mendling, J., Strembeck, M., Recker, J.: On the cognitive effectiveness of routing symbols in process modeling languages. In: Abramowicz, W. and Tolksdorf, R. (eds.) Business Information Systems. pp. 230–241. Springer Berlin Heidelberg (2010).
20. Liu, C., Li, Q., Zhao, X.: Challenges and opportunities in collaborative business process management: Overview of recent advances and introduction to the special issue. Inf. Syst. Front. 11, 201–209 (2009).
21. Mendling, J., Recker, J. C., Wolf, J.: Collaboration Features in Current BPM Tools. EMISA Forum. 32, 48–65 (2012).
22. Moody, D.: Cognitive Load Effects on End User Understanding of Conceptual Models: An Experimental Analysis. In: Benczur, A., Demetrovics, J., and Gottlob, G. (eds.) Advances in Databases and Information Systems. pp. 129–143. Springer Berlin Heidelberg (2004).
23. Scheer, A.-W., Nüttgens, M.: ARIS Architecture and Reference Models for Business Process Management. In: van der Aalst, W., Desel, J., and Oberweis, A. (eds.) Business Process Management. pp. 366–379. Springer Berlin Heidelberg (2000).
24. Moody, D. L.: Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions. Data Knowl. Eng. 55, 243–276 (2005).
25. Moreno-Montes de Oca, I., Snoeck, M., Reijers, H. A., Rodríguez-Morffi, A.: A systematic literature review of studies on business process modeling quality. Inf. Softw. Technol. 58, 187–205 (2015).
26. Wand, Y., Weber, R.: Research Commentary: Information Systems and Conceptual Modeling—A Research Agenda. Inf. Syst. Res. 13, 363–376 (2002).
27. Lindland, O. I., Sindre, G., Solvberg, A.: Understanding quality in conceptual modeling. IEEE Softw. 11, 42–49 (1994).
28. Nelson, H. J., Poels, G., Genero, M., Piattini, M.: A conceptual modeling quality framework. Softw. Qual. J. 20, 201–228 (2012).
29. Krogstie, J., Sindre, G., Jørgensen, H.: Process models representing knowledge for action: a revised quality framework. Eur. J. Inf. Syst. 15, 91–102 (2006).
30. Patig, S.: A practical guide to testing the understandability of notations. Conf. Res. Pract. Inf. Technol. Ser. 79, 49–58 (2008).
31. Moody, D. L., Sindre, G., Brasethvik, T., Sølvberg, A.: Evaluating the Quality of Process Models: Empirical Testing of a Quality Framework. In: Spaccapietra, S., March, S. T., and Kambayashi, Y. (eds.) Conceptual Modeling - ER 2002. pp. 380–396. Springer Berlin Heidelberg (2003).
32. Krogstie, J., Lindland, O. I., Sindre, G.: Defining quality aspects for conceptual models. In: Proceedings of the IFIP international working conference on Information system concepts: Towards a consolidation of views. pp. 216–231 (1995).
33. Webster, J., Watson, R. T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Q. 26, xiii – xxiii (2002).
34. vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In: ECIS 2009 Proceedings. pp. 2206–2217 (2009).
35. Cooper, H. M.: Organizing knowledge synthesis: a taxonomy of literature reviews. Knowl. Soc. 1, 104–126 (1988).
36. Netto, J. M., Franca, J. B. S., Baiao, F. A., Santoro, F. M.: A notation for Knowledge-Intensive Processes. In: Proceedings of the 2013 IEEE 17th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2013. pp. 190–195 (2013).

37. Dollmann, T., Houy, C., Fettke, P., Loos, P.: Collaborative business process modeling with CoMoMod: A toolkit for model integration in distributed cooperation environments. In: Proceedings of the 20th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE 2011. pp. 217–222 (2011).
38. Grosskopf, A., Edelman, J., Weske, M.: Tangible business process modeling - Methodology and experiment design. In: Rinderle-Ma, S., Sadiq, S., and Leymann, F. (eds.) Business Process Management Workshops. pp. 489–500. Springer Berlin Heidelberg (2010).
39. Schumm, D., Leymann, F., Streule, A.: Process Viewing Patterns. In: Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC. pp. 89–98 (2010).
40. Kathleen, N., Ross, B., Kriglstein, S.: Storyboard Augmentation of Process Model Grammars for Stakeholder Communication. In: IVAPP 2014: Proceedings of the 5th International Conference on Information Visualization Theory and Applications. pp. 114–121. , Lisbon (2014).
41. Allweyer, T., Schweitzer, S.: A Tool for Animating BPMN Token Flow. In: Mendling, J. and Weidlich, M. (eds.) Business Process Model and Notation. pp. 98–106. Springer Berlin Heidelberg (2012).
42. Rittgen, P.: Collaborative Business Process Modeling – Tool Support for Solving Typical Problems. In: International Conference on Information Resources Management (CONF-IRM). pp. 1–11 (2010).
43. Riemer, K., Holler, J., Indulska, M.: Collaborative process modelling-tool analysis and design implications. In: ECIS 2011 Proceedings (2011).
44. Ackermann, L., Schönig, S., Zeising, M., Jablonski, S.: Natural Language Generation for Declarative Process Models. In: Barjis, J., Pergl, R., and Babkin, E. (eds.) Enterprise and Organizational Modeling and Simulation. pp. 3–19. Springer International Publishing (2015).
45. Harman, J., Brown, R., Johnson, D., Rinderle-ma, S., Kannengiesser, U.: Virtual Business Role-Play : Leveraging Familiar Environments to Prime Stakeholder Memory During Process Elicitation. 1, 166–180.
46. Bernstein, V., Soffer, P.: How Does It Look? Exploring Meaningful Layout Features of Process Models. In: Persson, A. and Stirna, J. (eds.) Advanced Information Systems Engineering Workshops. pp. 81–86. Springer International Publishing (2015).
47. Khlif, W., Ben-Abdallah, H.: Integrating semantics and structural information for BPMN model refactoring. In: Computer and Information Science (ICIS) (2015).
48. Berstein, V., Soffer, P.: Identifying and Quantifying Visual Layout Features of Business Process Models. In: Gaaloul, K., Schmidt, R., Nurcan, S., Guerreiro, S., and Ma, Q. (eds.) Enterprise, Business-Process and Information Systems Modeling. pp. 200–2013. Springer International Publishing (2015).
49. Chinces, D., Salomie, I.: Optimizing spaghetti process models. In: International Conference on Control Systems and Science. pp. 506–511 (2015).
50. Kannengiesser, U., Oppl, S.: Business Processes to Touch : Engaging Domain Experts in Process Modelling.
51. Hipp, M., Strauss, A., Michelberger, B., Mutschler, B.: Enabling a User-Friendly Visualization of Business Process Models. In: International Conference on Business Process Management. pp. 1–12. Springer International Publishing.
52. Petrusel, R., Mendling, J., Reijers, H.A.: Task-specific visual cues for improving process model understanding. Inf. Softw. Technol. 79, 63–78 (2016).
53. Oppl, S.: Articulation of work process models for organizational alignment and informed information system design. Inf. Manag. 53, 591–608 (2016).
54. Geryville, H. M., Bouras, A., Ouzrout, Y., Sapidis, N. S.: Collaborative Product and Process Model : Multiple Viewpoints Approach. In: Technology Management Conference (ICE) (2006).

55. Hanser, M., Ciccio, C. Di, Mendling, J.: A Novel Framework for Visualizing Declarative Process Models. In: ZEUS Workshop. pp. 27–28 (2016).
56. Emens, R., Vanderfeesten, I., Reijers, H. A.: The Dynamic Visualization of Business Process Models: A Prototype and Evaluation. In: Proceedings of the 4th International Workshop on Theory and Application of Visualizations and Human-centric Aspects in Processes (2015).
57. De Smedt, J., De Weerdt, J., Serral, E., Vanthienen, J.: Improving Understandability of Declarative Process Models by Revealing Hidden. In: CAiSE. pp. 83–98 (2016).
58. Weber, B., Reichert, M., Mendling, J., Reijers, H. A.: Refactoring large process model repositories. Comput. Ind. 62, 467–486 (2011).
59. Reijers, H. A., Freytag, T., Mendling, J., Eckleder, A.: Syntax highlighting in business process models. Decis. Support Syst. 51, 339–349 (2011).
60. Antunes, P., Simões, D., Carriço, L., Pino, J. A.: An end-user approach to business process modeling. J. Netw. Comput. Appl. 36, 1466–1479 (2013).
61. Kuester, J., Voelzer, H., Favre, C., Branco, M., Czarnecki, K.: Supporting Different Process Views through a Shared ProcessModel. In: Van Grop, P., Ritter, T., and Rose, L. M. (eds.) Modelling Foundations and Applications. pp. 20–36. Springer Berlin Heidelberg (2013).
62. Mafazi, S., Mayer, W., Grossmann, G., Stumptner, M.: A Knowledge-based Approach to the Configuration of Business Process Model Abstractions. In: 1st International Workshop on Knowledge-intensive Business Processes. pp. 60–74 (2012).
63. Polyvyanyy, A., Smirnov, S., Weske, M.: Process model abstraction: A slider approach. In: Proceedings - 12th IEEE International Enterprise Distributed Object Computing Conference, EDOC 2008. pp. 325–331 (2008).
64. Meyer, A., Weske, M.: Data Support in Process Model Abstraction. In: Conceptual Modeling. pp. 292–306. Springer Berlin Heidelberg (2012).
65. Fernández-Ropero, M., Pérez-Castillo, R., Piattini, M.: Graph-Based Business Process Model Refactoring. Simpda. 16–30 (2013).
66. Kolb, J., Reichert, M., Weber, B.: Using Concurrent Task Trees for Stakeholder-centered Modeling and Visualization of Business Processes. In: Oppl, S. and Fleischmann, A. (eds.) S-BPM ONE - Education and Industrial Developments. pp. 237–251. Springer Berlin Heidelberg (2012).
67. Smirnov, S., Reijers, H. A., Weske, M.: A Semantic Approach for Business Process Model Abstraction. In: Mouratidis, H. and Rolland, C. (eds.) Advanced Information Systems Engineering. pp. 497–511. Springer Berlin Heidelberg (2011).
68. Thom, L. H., de Oliveira, J. P. M., Gassen, J. B., Abel, M.: Towards an Ontological Process Modeling Approach. In: ONTOBRAS-MOST. pp. 242–247 (2012).
69. Fleischmann, A., Metasonic AG, Kannengiesser, U., Schmidt, W., Stary, C.: Subject-Oriented Modeling and Execution of Multi-agent Business Processes. Proc. - 2013 IEEE/WIC/ACM Int. Conf. Intell. Agent Technol. IAT 2013. 2, 138–145 (2013).
70. Charfi, A., Müller, H., Mezini, M.: Aspect-Oriented Business Process Modeling with AO4BPMN Modelling Foundations and Applications. In: Kühne, T., Selic, B., Gervais, M.-P., and Terrier, F. (eds.) Modelling Foundations and Applications. pp. 48–61. Springer Berlin Heidelberg (2010).
71. Heitkötter, H.: A Framework for Creating Domain-specific Process Modeling Languages. In: Proceedings of the 7th International Conference on Software Paradigm Trends (ICSOFT). pp. 127–136 (2012).
72. Schnabel, F., Gorronogoitia, Y., Radzimski, M., Lecue, F., Mehandjiev, N., Ripa, G., Abels, S., Blood, S., Mos, A., Junghans, M., Agarwal, S., Vogel, J.: Empowering business users to model and execute business processes. In: zur Muehlen, M. and Su, J. (eds.) Business Process Management Workshops. pp. 433–448. Springer Berlin Heidelberg (2011).

73. Natschläger, C.: Deontic BPMN. In: Hameurlain, A., Liddle, S. W., Schewe, K.-D., and Zhou, X. (eds.) Database and Expert Systems Applications. pp. 264–278. Springer Berlin Heidelberg (2011).
74. Fahland, D., Weidlich, M.: Scenario-based process modeling with GRETA. In: Proceedings of the Business Process Management 2010. pp. 52–57 (2010).
75. Caetano, A., Zacarias, M., Silva, A. R., Tribolet, J.: A Role-Based Framework for Business Process Modeling. In: Proceedings of the 38th Hawaii International Conference on Systems Sciences (HICSS'05). pp. 1–7 (2005).
76. Born, M., Kirchner, J., Müller, J. P.: Context-driven business process modelling. In: Joint Proceedings of the 4th International Workshop on Technologies for Context-Aware Business Process Management, TCoB 2009. AT4WS 2009. AER 2009. MDMD 2009. In Conjunction with ICEIS 2009. pp. 17–26 (2009).
77. Cappelli, C., Santoro, F. M., Cesar Sampaio Do Prado Leite, J., Batista, T., Luisa Medeiros, A., Romeiro, C. S.: Reflections on the modularity of business process models: The case for introducing the aspect-oriented paradigm. Bus. Process Manag. J. 16, 662–687 (2010).
78. Krumnow, S., Decker, G.: A Concept for Spreadsheet-Based Process Modeling. In: Mendling, J., Weidlich, M., and Weske, M. (eds.) Business Process Modeling Notation. pp. 63–77. Springer Berlin Heidelberg (2010).
79. Dyke, N. W. Van: Generating Hypertext Explanations for Visual Languages. In: Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia: Links, Objects, Time and Space - Structure in Hypermedia Systems. pp. 301–302 (1998).
80. Reichert, M.: Visualizing Large Business Process Models: Challenges, Techniques, Applications. In: La Rosa, M. and Soffer, P. (eds.) Business Process Management Workshops. pp. 725–736. Springer Berlin Heidelberg (2013).
81. Bittmann, S., Metzger, D., Fellmann, M., Thomas, O.: Additional Information in Business Processes: A Pattern-Based Integration of Natural Language Artefacts. In: Modellierung. pp. 137–152 (2014).
82. Kolb, J., Leopold, H., Mendling, J., Reichert, M.: Creating and Updating Personalized and Verbalized Business Process Descriptions. In: Grabis, J., Kirikova, M., Zdravkovic, J., and Stirna, J. (eds.) The Practice of Enterprise Modeling. pp. 191–205. Springer Berlin Heidelberg (2013).
83. Betz, S., Eichhorn, D., Hickl, S., Klink, S., Koschmider, A., Li, Y., Oberweis, A., Trunko, R.: 3D Representation of Business Process Models. MobIS. 73–87 (2008).
84. Polyvyanyy, A., Smirnov, S., Weske, M.: Reducing complexity of large EPCs. In: Modellierung betrieblicher Informationssysteme: Modellierung zwischen SOA und Compliance Management. pp. 195–207 (2008).
85. Brown, R., Recker, J., West, S.: Using virtual worlds for collaborative business process modeling. Bus. Process Manag. J. 17, 546–564 (2011).
86. Effinger, P.: A 3D-Navigator for Business Process Models. In: La Rosa, M. and Soffer, P. (eds.) Business Process Management Workshops. pp. 737–743. Springer Berlin Heidelberg (2013).
87. Effinger, P.: Layout Patterns with BPMN Semantics. In: Dijkman, R., Hofstetter, J., and Koehler, J. (eds.) Business Process Model and Notation. pp. 130–135. Springer Berlin Heidelberg (2011).
88. Effinger, P., Siebenhaller, M., Kaufmann, M.: An Interactive Layout Tool for BPMN. In: IEEE Conference on Commerce and Enterprise Computing. pp. 399–406 (2009).
89. Shen, M., Liu, D.-R.: Coordinating Interorganizational Workflows Based on Process-Views. In: Mayr, H. C., Lazansky, J., Quirchmayr, G., and Vogel, P. (eds.) Database and Expert Systems Applications. pp. 274–283. Springer Berlin Heidelberg (2001).

A Meta Model Based Extension of BPMN 2.0 for Mobile Context Sensitive Business Processes and Applications

Julian Dörndorfer[1] and Christian Seel[1]

[1] Hochschule Landshut, Computer Science, Landshut, Germany
{julian.doerndorfer,christian.seel}@haw-landshut.de

Abstract. Smart devices like smartphones or tablets have become ubiquitous, which affected many daily work activities like maintaining contacts via a mobile CRM anywhere, anytime. Thus, business processes can now be executed independently of an employee's location. In addition, mobile devices have the possibility to measure physical quantities through sensors, like location or acceleration. Moreover, the connection to wireless networks made it possible to query context information like customer history. These context information can be used to adapt mobile business processes and the mobile application that support them. But in order to use this advantage, mobile sensor data has to be reflected in the business process model. As current languages for process aware information systems, such as BPMN, do not support the influence of mobile context information, we propose an extension of the BPMN that will enable the modeling of mobile context sensitive business processes.

Keywords: Business Process Modeling, Context, Mobile Business Processes, DSML, BPMN Extension

1 The Impact of Context on Mobile Business Processes

The modeling and implementation of business processes is a standard approach in Information Systems theory and practice [1–5]. In the last few years, mobile devices like smart phones, tablets, wearables, etc., affected both the work of employees and employers and thus the modeling and conduction of business processes. In a survey by Intel Research [6], 47% of employees stated that they used smart phones for their daily tasks in the workplace. 18% of the respondents use tablets for their daily job. New issues like bringing your own device [7, 8] and mobile device management [9] illustrate the impact of using mobile devices in the workplace. For field worker, who visit customers, mobile device are even more common. This change from stationary computers to mobile devices also concerns business processes. Processes, which are supported by information systems became nomadic and can now be executed anywhere and anytime. However, not only the location of the execution has changed, but the sensors of mobile devices can also deliver valuable data that affect business processes. These data can be used to directly adapt business processes to the current context that is detected by sensors. Current sensors usually measure location, acceleration, brightness and, depending on the device, other physical quantities. In addition to

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Dörndorfer, J.; Seel, C. (2017): A Meta Model Based Extension of BPMN 2.0 for Mobile Context Sensitive Business Processes and Applications, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 301-315

connecting to the mobile or wireless network, they can query data from different information systems or other machines [10]. Therefore, mobile devices are capable of measuring context and evaluating it via algorithms. Thereby, DEY defines ‘context’ as any information that can be used to characterize the situation of an entity [11]. An entity can be a person, place or object which is considered relevant to the interaction between a person and an application. Context can be used to pre-select services to reduce the information overload for a user [12].

The various possibilities of mobile devices can support the user by making decisions in a business process or when carrying out a task. FALK and LEIST [13] were able to show that mobile applications have positive impact on business processes, like cost and time reduction, as well as quality and flexibility improvements. Furthermore, applications which measure and interpret the context can automatically execute parts of the business process without the involvement of any other party. Additionally, they can adapt the process, for example skip or block [14] an activity, and influence the execution. A good example are the employment protection inspectors. From time to time they have to control the working condition of the employees, the correct montage or installation of machines and their proper functionality. In some countries like Germany these regular inspections are prescribed by law like the ‘Betriebssicherheitsverordnung’ [15]. The inspectors have to check and to document different things depending on their location, time or presence of employees. If the inspectors are in an assembly hall, they could check the machines or, if the machines are running, they could control the work clothes. If the assembly hall has security areas, they could control the security mechanisms and precautions for the employees. All these tasks are mainly depending on context information, which can easily be accessed by smart devices. They also support the task executions by showing what to control and by documenting completed checkups. All these process steps are immediately interrupted if the context switches to emergency because of an accident.

This mobile context sensitive business process gives a good impression of how context data can be used to support and ease the work of employees and increase the efficiency of business processes. However, to plan mobile context sensitive business processes a modeling language has to support context. The standard languages in the area of information systems, like Business Process Model and Notation (BPMN), Event-driven Process Chain (EPC) or the Unified Modeling Language (UML), do not support the modeling of context [16–18]. However, some approaches were made to extend the standard modeling languages to improve the configuration and flexibility of them (cf. section 2), these approaches where neither designed for mobile business processes nor mobile sensor data.

As BPMN 2.0 [19] became the de-facto standard [20] and allows to automate business processes based on execution engines, it seems reasonable to extend BPMN with adaption to mobile context. Hence, the following research question (RQ) arise:

RQ.1: *To what extend could the standard elements of the BPMN 2.0 specification be used to model contextual influences?*

RQ.2: *How could extension of the BPMN be structured to ensure an appropriate modeling of mobile context sensitive business processes?*

The remainder of this paper is structured as follows: Section 2 introduces a brief overview about existing approaches and related work in this field. In the beginning of section 3 the existing BPMN elements are analyzed for the capability to express context (**RQ.1**). Thereafter, the requirements for the extension are discussed and collected. Section 4.2 presents a meta-model based BPMN extension (**RQ.2**). Afterwards an example model is shown in order to prove the applicability of the BPMN extension. The paper ends with a conclusion and outlook to further research in context sensitive mobile applications.

2 Related Work and Research Method

The term context was defined among others [21–23] by DEY [24]. DEY'S definition is well known and accepted in the scientific community. He also declares an application as context-aware "if it uses context to provide relevant information and/or services to user, where relevancy depends on the user's task" [24].

Several approaches have been conducted to integrate context into business processes. ROSEMANN et al. [25] claim that modeling languages have to be more flexible to model context. Further they state that an increased attention on flexibility took place in the research area, which leads to a decreasing time-to-market for products [26]. Therefore, the result is a demand for higher process flexibility [27]. In particular, ROSEMANN et al. show the limitation of the actual EPC language and the lack of supporting context modeling. They also present a framework that helps to understand the different types of context and their impact on business processes. In addition, they categorize the different influence factors of a business process. However, they do not present an appropriate way to integrate the identified context in a business process. C-EPC is an extension of EPC to make it more configurable for decisions at runtime [28]. In [29] an approach to identify all variants of a business process depending on its context is presented. LA VARA et al. aim to reveal all possibilities of a business process and integrate them into one model. The outcome is a large and complex model. A way to identify and apply context on business processes is introduced in [30]. However, it is more of a theoretical framework to identify context. But the authors do not show how a context sensitive business process could be designed.

Furthermore two extensions for UML are published by AL-ALSHUHAI and SIEWE. In [31] they are extending the class diagram with additional annotations. The second paper [32] expands the activity diagram to mark context sensitive areas or sequences.

HEINRICH and SCHÖN [33] mention that business processes must consider „non-static" context events which change the process conduction, like an upcoming thunderstorm in an outdoor process. They further present an algorithm which supports automated process planning for context-aware processes, but no modelling representation in BPMN 2.0. CONFORTI et al. [34] presented an approach to cope with process risks based on sensor evaluation. Furthermore, they presented a way to model these risks and when they occur. However, they see sensors only as a source for risk evaluation, but context and its evaluation via sensors is more than only risk analyzation for business processes.

All of the presented papers do not consider how to model mobile context sensitive business processes. Therefore, this paper aims to present a new context extension for BPMN and creates a new artifact based on the design research approach by HEVNER et al. in [35]. Therefore, the meta-model of the extension will be created with Meta Object Facility (MOF) 2.0 [36]. New notation will be developed and a guideline to assist the user presented.

3 Requirements on Context Modeling of Mobile Business Processes

3.1 Analysis of the Existing BPMN Elements

To answer the **RQ.1** the BPMN 2.0 core elements will be examined in this section. They offer a few possibilities to model contextual mobile business processes. These possibilities consist of all the existing gateways (exclusive, inclusive, complex etc.) or intermediate events. Modeling context in mobile business processes using only gateways leads to a large and confusing model (cf. figure 1). Intermediate boundary events, which could be attached on activities, would be another reasonable option to model context influences. However, no event exists that really matches the description of context. For example the intermediate event *Message* "can be used to either send a Message or receive a Message" [18]. It can be used to symbolize that a context change "sends" an information, which leads to a reaction. At a first glance this looks like a promising method to model context, but a closer look reveals the problem. The event *Message* is, like other alternative events, very generic and could have several meanings. In addition, there is no possibility to describe a proper context expression. The result of using this event for depicting context influences would be a large documentation, which is counterproductive to the aims of a process visualization. Moreover, existing BPMN notation cannot sufficiently cope with non-static context [33], like changing parameters at runtime. A mix of intermediate events and gateways seems to be even worse. Confusing models with a documentation is not the idea of business process modeling. The result of the analysis is, that the existing elements of BPMN cannot be used to model context.

However, the specification of the BPMN offers a possibility to extend existing elements or add new elements to it [18]. To develop an extension of the BPMN, the requirements have to be imposed and discussed.

3.2 Requirements for the Extension

As described before, no explicit symbol or label exists which marks that an activity is influenced by context. Furthermore, not just one activity but a whole process sequence could depend on context data. In the example from section 1 an employment protection inspector has been introduced. She or he controls the working and security conditions at the workplace and inspects machines.

Figure 1 depicts a simplistic inspection process. Depending on her or his location (office building or different assembly halls) and present workers, she or he controls the first-aid kit or the security mechanisms. However, it is difficult to see that an activity or sequence depends on context information. Many decisions have to be made and modeled by gateways, which leads to large and complex models. Thus, the first requirement for the extension is an element which marks an activity or a process sequence depending on the context.

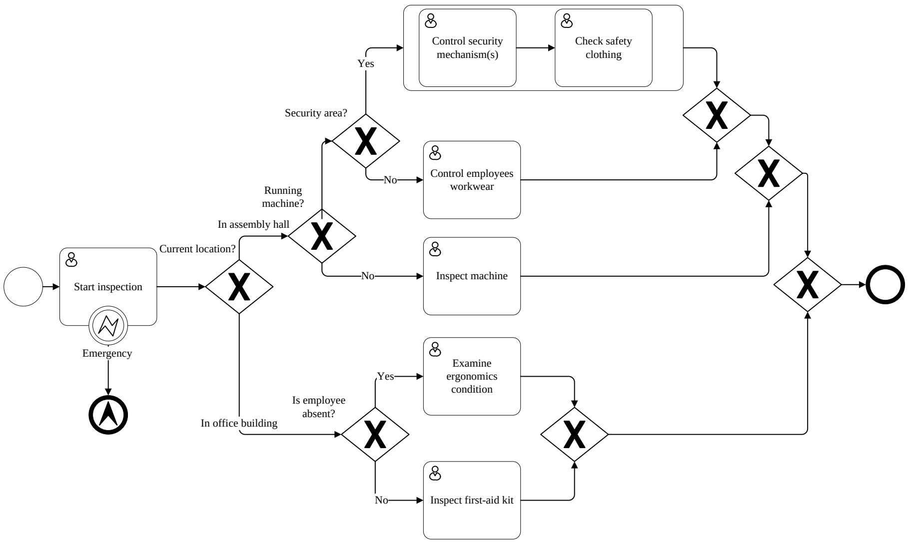


Figure 1. “Employment protection process” with BPMN 2.0 core elements

BPMN events are divided into interrupting and non-interrupting. Also a changing context can be an interrupting or non-interrupting event for an activity or sequence. For instance, if an emergency happens in the assembly hall, the inspector would take suitable measures and she or he would interrupt his or her current activity. At the moment, it is hard to model such an occurrence, because every activity in the process would have an attached exception event, like the one depicted in figure 2.

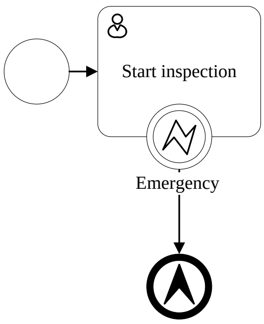


Figure 2 – Example for exception modeling in traditional BPMN 2.0

This would inflate the model unnecessarily. However, not every context change leads to an interruption of the current activity or sequence. Let us assume that a shift change happens during the first-aid kit check. Nevertheless, the context event “shift change” does not interrupt the activity “inspect the first-aid kit”. Therefore it is required that context events can be interrupting or non-interrupting. Another requirement is to express context dependencies in a short but clear way. A decision can depend on more than one context data. A textual description could lead to long sentences or even to an extra documentation file. To avoid this, a brief description method should be developed and applied.

The time to evaluate the context is important. Depending on the concrete business process, two possibilities exist. First, the context measurement is necessary at one explicit point in time. The evaluation of the context “location” only needs to be queried at the beginning of the inspection, because the location remains unchanged. In other processes the context information will be evaluated continuously. The context information “emergency” has to be queried in a tight frequency (e.g. 2 seconds) to show supportive information on the smart devices. It would be impractical to query this information only once at the beginning. Therefore, it must be possible to express the time of evaluation in the mobile business model to cope with non-static context.

In section 1 the advantages of using a standard modeling language have been discussed. To utilize these advantages the extension has to be compatible to the BPMN 2.0 standard. Thus, this is an additional requirement for the extension.

Context is measured through sensors. We define sensors not only as devices that detect and respond to a physical in- or output, like a hygrometer or a temperature sensor. It could also be a database or an application from which information could be requested or even any machine that is accessible via a network connection. In addition, a context information could be based upon different sensors. For example the context information “weather” is based on the following sensors: “temperature”, “humidity”, “condensation”, etc. Furthermore, the context information “weather” could be a “sensor” for the context “traffic”, in order to warn of snow in a traffic information system. So, context and its sensors could have interdependencies. This shows that context information is hard to model due to its complexity. Therefore, we decided to split the sensor modeling from the BPMN model and introduce an extra model. A link will be established to connect these two models. All the discussed requirements of this section are summarized in Table 1.

Table 1 - Requirements for the extension

Requirement	Description	Model
Context marker	A way to mark an activity or a sequence as context depending.	BPMN extension
Context event is interrupting / non-interrupting	A context event can be interrupting or non-interrupting.	BPMN extension

Requirement	Description	Model
Context expression language	Express a context condition in a brief way.	BPMN extension
Evaluation point	The point in time at which the context information should be evaluated.	BPMN extension
Compatibility to BPMN 2.0	The extension has to be compatible to BPMN 2.0.	BPMN extension
Sensor modeling	A possibility to model the dependencies of a context information to sensors or other context information.	Dedicated DSML for sensor modeling

4 Meta Model Based Extension

4.1 The BPMN Extensibility

A common way to extend modeling languages is a meta model extension [37]. A meta model is besides a notation and a guideline ([38] p. 36) [39] the first step to develop a complete modeling technique. A meta model is a model (M2) which describes a language (L1) to create a model (M1) of an object [40]. BPMN describes all of its elements in a class diagram, which is the meta-language for it. In addition, the BPMN specification itself describes a mechanism to extend the model to the needs of domain ([18] p. 57). The mechanism ensures the core validity and enables a way to integrate domain specific concepts. It consist mainly of four classes: Extension, ExtensionDefinition, ExtensionAttributeDefinition, ExtensionAttributeValue. These classes allow to expand existing elements through additional attributes or to enhance the BPMN meta model. To meet the compatibility requirement to BPMN 2.0 we use this well described mechanism to extend the BPMN model.

4.2 Meta Model of the Context Extension

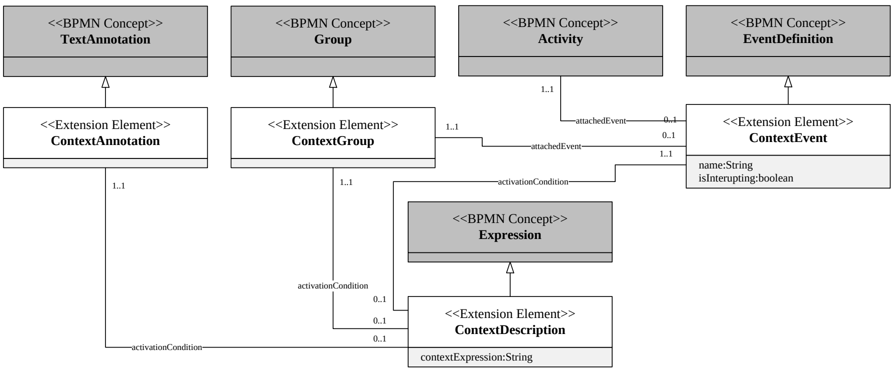


Figure 3. Meta model of Context4BPMN

Based on the in section 3.2 described requirements a meta model for Context4BPMN has been developed. The result is depicted in figure 3. Object Management Group's (OMG) standard BPMN classes are colored in grey. To mark context sensitive areas in a BPMN the Group class is extended by the *ContextGroup* class. The *ContextEvent* class creates the new context event which can be interrupting and non-interrupting as well as attached to an activity or a *ContextGroup*. Whether it is interrupting or not, is represented by the boolean variable. The *ContextDescription* class inherits from the *Expression* class to allow expressing the activation conditions of an activity or sequence. A *ContextDescription* includes the *contextExpression* to define the activation conditions, which are stored in a string. The grammar of the expression is explained in section 4.3. The *ContextDescription* could be used by the *ContextAnnotation*, *ContextMarker* and *ContextEvent*.

4.3 Syntax of the Context Extension and Example

In figure 4 the new element's graphical notations of the Context4BPMN extension are presented. The first two elements are the *Intermediate Context Events*, which add two new event types to the *Event* elements. The first element with the two solid circles and the eye is the *Interrupting Context Event*. Thereafter comes the *Non-Interrupting Context Event*. Both can be attached on *Activities* or *ContextGroups* concluded from the meta model. The *Context Annotation* is an enhancement of the *Text Annotation*. A little icon in the upper inner corner marks this new element. It can be attached to nearly every element like the "original" *Text Annotation*. The last element is the *Context Group,* which marks a activity or a sequence as context depending.

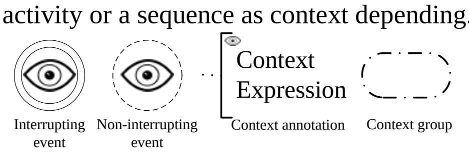


Figure 4. The new notations

```
<ContextExpression>::= [<EvaluationPoint>";"]<ContextTerm>
<ContextTerm>::= <Variable> <Comparison> <Value> |
  <ContextTerm> <LogicOperator> <ContextTerm>
<Comparison>::= "=" | "!=" | "<=" | ">=" | "<" | ">"
<LogicOperator>::= "AND" | "OR" | "XOR"
<EvaluationPoint>::= "Evaluation Point = continuous @"
  <Interval>| "Evaluation Point = " <Time>
<Time>::= -->"TimeDefintion according to RFC 5322"
<Interval>::= -->"IntegerNumber" "msec" | "sec" |"min" |
  "hours" | "days"
<Variable>::= -->"StringIdentifier in UTF-8"
<Value>::= -->"StringIdentifier in UTF-8"
```

Figure 5. EBNF Grammar for context expressions

In order to express context in a structured way that can be evaluated automatically a context expression language is developed. The context free grammar for this language has been defined in the extended Backus–Naur Form (EBNF) [41] in figure 5. To shorten some basic definitions, like integer numbers or date, we link with "-->" to standards, like the formal definition in RFC 5322 for date formats. A context expression contains a context term, which basically consists of "variable" "comparator" "value". The "variable" is the name of the context, like "location" or "status". It has to be unique in a business process and is also the link to the sensor model declaring how to measure it. Examples of the language are presented in figure 8.

4.4 Guideline for the Extension

To get a complete modeling technique a guideline for the modeler to apply the language extension is necessary ([38] p. 36) [39]. Three initial situations of mobile context sensitive modeling exists. First, a new process has to be modeled from scratch. Second, an existing process has to be remodeled as context sensitive. The last situation is when different variations of a business process depending on context variables exist and have to be merged into one business model. The first two situations differ only in the beginning, when a new business process has to define its activities, whereas an existing process has to be decomposed. Hereafter, the next steps are identical. Figure 6 and figure 7 are depicting the guideline for the different situations, which are modeled in BPMN. The guideline for the first and the second situation are an expanded version from the "procedure for context identification" by ROSEMANN et al. [25]. If a process exists, a decomposing would be necessary to get a list of activities, whereas for a new process a procedure has to be defined. At the end of both tasks a list of activities would be the outcome. The next step would be to determine context influences. The aim of this step is to identify the relevant context, which will be used in the next step to detect the context depending activities. Hereafter, a mobile context sensitive business process model could be developed with the identified context depending activities.

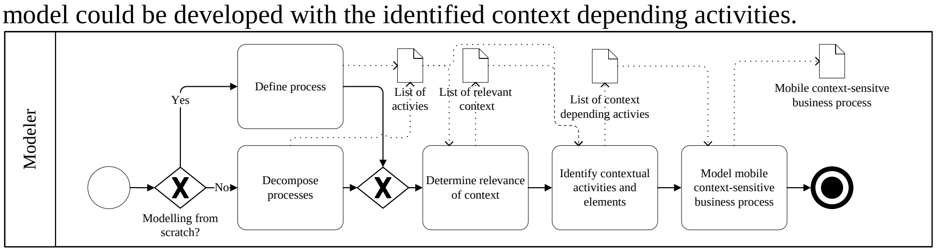


Figure 6. Guideline for initial situation one and two

If different variations of the same business process exist, the first task has to determine all variations. The output, which is a list of a variations, will be used to identify the relevant context, causing the variations. The third step is to merge all alternatives into one business model. For this task model matching algorithms can be used like the one proposed in [42]. Identical activities or sequences are context independent, whereas deviations in the models are most likely context depending. This holistic model will be used to identify the context depending activities or sequences.

These activities together with the relevant context variables will be used to model the context sensitive business process.

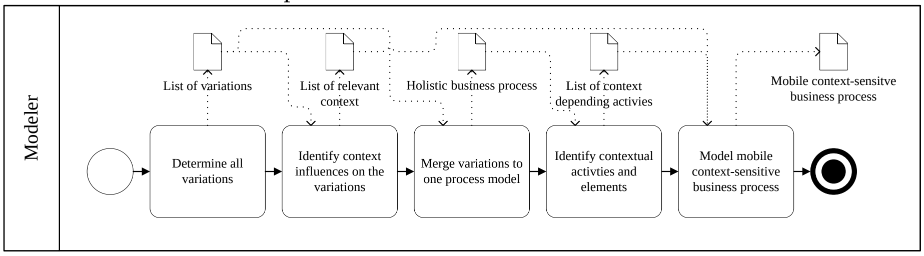


Figure 7. Guideline for the third initial situation

5 Discussion and Further Research

5.1 Requirement Examination

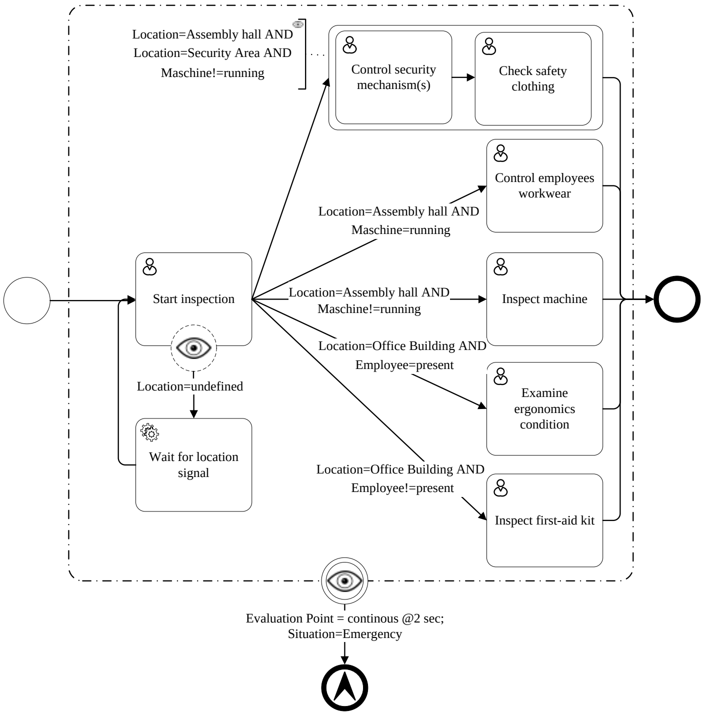


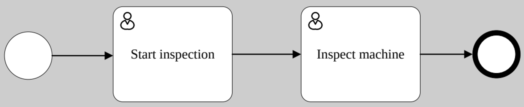
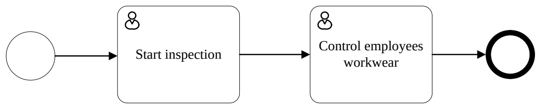
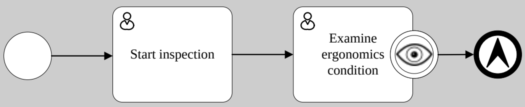
Figure 8. Inspection example depicted with the extension Context4BPMN

As stated in the beginning of this paper, it was intended to develop a possibility to model mobile context sensitive business processes. Therefore, we aimed to create an extension for BPMN to save effort and time in developing and establishing a new domain language profiting from well-known mechanisms and tool support. A brief evaluation of the created artifact as recommended by HEVNER et al. [35] will be presented. Figure 8 again depicts the example “employment protection” from section 3.2, but utilizing the new elements. The whole sequence is context depending, therefore the context group marks all depicted activities. Attached to it is an interrupting context event, which will be triggered if the situation changes from normal to emergency (non-static context). The situation could be changed at once, therefore the information is requested every 2 seconds. It would be displayed in the supporting mobile application. Further steps in the process would be defined in the catching escalation event. The “Start inspection” activity now has the non-interrupting context event attached to define the task if no location could be determined by the application.

Afterwards, one activity depending on the evaluated context will be selected. This can be seen as selecting an activity out of a pool of activities at runtime. The selection will be done by evaluating the context expression. These expressions are shown on the sequence flows and in the context annotation, attached to the brief sub-process.

Table 2 depicts the applied process steps for a certain context situation. For instance, the first row shows the process when the inspector is in the assembly hall, the machine is not running and no emergency interferes the inspection.

Table 2. Different versions of the process at runtime

Variation	Context	Applied version of the process
1	Assembly hall && machine not running && no emergency	
2	Assembly hall && machine running && no emergency	
3	Office building && employee present && emergency	

In section 3.2 a table with the requirements for the extension was introduced. To reconcile the requirements with the developed extension features we conducted a comparison.

Table 3. Comparison between the requirements and the features of Context4BPMN extension

Requirement	Feature	Match
Context marker	Context Group	✓
Context event is interrupting / non-interrupting	Intermediate Context Event	✓

Requirement	Feature	Match
Context expression language	Context expressions could be placed on flows and in Context Annotations	✓
Evaluation point	The Evaluation point can be stated in the context expression language	✓
Compatibility to BPMN 2.0	The extension uses the BPMN enhancement mechanism to secure the compatibility to the 2.0 version	✓
Sensor modeling	Due to lack of space for this paper the sensor model could not be presented. Only the link to it is established. Therefore, this feature is not completely described	-

As can be seen from table 3 all direct requirements to the BPMN extension are fulfilled and depicted in figure 8. A context marker is established as a context group element in the extension. Even though two intermediate context events were added, one interrupting and one non-interrupting, context expressions can now be stated on flow elements or on context annotations. It is also possible to express an evaluation time in the context expression language. A link to the sensor model is as well established. Hence, it can be claimed that all stated requirements are fulfilled.

5.2 Outlook

The main contribution of this article is to provide an extension to model mobile context sensitive business processes. It provides the possibility to model engineers to plan such processes in a precise, detailed and comprehensive way. It also enables to model the influence of context, like intermediate events on groups, which has not been available before.

There are some tasks for further research in this area: First of all, a way to model the relationships between context and sensors, furthermore, to model the interdependencies between two pieces of context information. The reason to prefer the creation of a new language instead of integrating it into a standard language like BPMN is that at firstly the model would become unnecessarily big and complex. Secondly, the inter-dependencies of context information and measuring by sensors is not a necessary part of a business process language.

Modeling a business process is just the first step in the business process lifecycle (BPL) [43]. The implementation, execution and controlling are the remaining steps in the lifecycle. Since mobile context sensitive business processes obviously need to measure context and are supported by an application on smart devices, an automated or semi-automated way to generate code from the business process would be helpful to increase the fluency between modeling and implementation phase. The logical context expressions can be used to generate decisions in the application program. Furthermore, the sensor model can be utilized to pre-generate classes and interfaces. In an additional step the use of the gathered context data from the execution of a business process will be investigated. They could be used to identify problems in the execution and therefore interesting for the controlling phase.

References

1. Scheer, A.-W.: ARIS - business process modeling. Springer, Berlin u.a. (2000)
2. Bichler, M., Frank, U., Avison, D., Malaurent, J., Fettke, P., Hovorka, D., Krämer, J., Schnurr, D., Müller, B., Suhl, L., et al.: Erratum to. Theories in Business and Information Systems Engineering. Bus Inf Syst Eng (2016)
3. Vom Brocke, J., Rosemann, M. (eds.): Handbook on Business Process Management 2. Strategic Alignment, Governance, People and Culture. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg (2010)
4. Hammer, M., Champy, J.: Reengineering the corporation. A manifesto for business revolution. Harper Business, New York, NY (1993)
5. Becker, J., Kugeler, M., Rosemann, M. (eds.): Process management. A guide for the design of business processes. Springer, Berlin (2011)
6. Intel IT Center: Mobile Computing Trends: Insight into Today's Workforce (2013)
7. Morabito, V. (ed.): Trends and challenges in digital business innovation. Springer, Cham (2014)
8. Kerr, D., Koch, C.: A Creative and Useful Tension? Large Companies Using "Bring Your Own Device". In: Bergvall-Kåreborn, B., Nielsen, P.A. (eds.) Creating Value for All Through IT. IFIP WG 8.6 International Conference on Transfer and Diffusion of IT, TDIT 2014, Aalborg, Denmark, June 2-4, 2014. Proceedings, 429, pp. 166–178. Springer Berlin Heidelberg, Berlin, Heidelberg, s.l. (2014)
9. Rhee, K., Eun, S.-K., Joo, M.-R., Jeong, J., Won, D.: High-Level Design for a Secure Mobile Device Management System. In: Marinos, L., Askoxylakis, I. (eds.) Human aspects of information security, privacy, and trust. First international conference, HAS 2013, held as part of HCI International 2013, Las Vegas, NV, USA, July 21 - 26, 2013 ; proceedings, 8030, pp. 348–356. Springer, Berlin (2013)
10. Fortino, G., Trunfio, P.: Internet of things based on smart objects. Technology, middleware and applications (2014)
11. Dey, A.K.: Understanding and Using Context. Personal and Ubiquitous Computing 5, 4–7 (2001)
12. Heinrich, B., Lewerenz, L.: A Novel Concept for the Usage of Mobile Information Services. In: Linnhoff-Popien, C., Zaddach, M., Grahl, A. (eds.) Marktplätze im Umbruch. Digitale Strategien für Services im mobilen Internet, pp. 319–329. Springer Vieweg, Berlin (2015)
13. Falk, T., Leist, S.: Effects of mobile solutions for improving business processes. ECIS 2014 Proceedings (2014)
14. Gottschalk, F., van der Aalst, W.M.P., Jansen-Vullers, M.H.: Configurable Process Models — A Foundational Approach. In: Becker, J., Delfmann, P. (eds.) Reference Modeling. Efficient Information Systems Design Through Reuse of Information Models, pp. 59–77. Physica-Verlag, Heidelberg (2007)
15. Verordnung über Sicherheit und Gesundheitsschutz bei der Verwendung von Arbeitsmitteln (Betriebssicherheitsverordnung). BetrSichV (2016)

16. Object Management Group (OMG): Unified Modeling Language (2015)
17. Keller, G., Nüttgens, M., Scheer, A.-W.: Semantische Prozeßmodellierung Semantische Prozeßmodellierung auf der Grundlage „Ereignisgesteuerter Prozeßketten (EPK)“. Saarbrücken (1992)
18. Object Management Group (OMG): Business Process Model and Notation (BPMN), Version 2.0 (2011)
19. International organization for standardization (iso): Information technology. Object Management Group Business Process Model and Notation (2013)
20. Braun, R., Schlieter, H., Burwitz, M., Esswein, W.: Extending a Business Process Modeling Language for Domain-Specific Adaptation in Healthcare. In: Wirtschaftsinformatik Proceedings 2015, pp. 468–481 (2015)
21. Weiser, M.: The Computer for the 21st Century. Sci Am 265, 94–104 (1991)
22. Schmidt, A., Beigl, M., Gellersen, H.-W.: There is more to context than location. Computers & Graphics 23, 893–901 (1999)
23. Schilit, B., Adams, N., Want, R.: Context-Aware Computing Applications. In: First Workshop on Mobile Computing Systems and Applications (WMCSA), pp. 85–90 (1994)
24. Dey, A.K.: Providing Architectural Support for Building Context-aware Applications. Georgia Institute of Technology, Atlanta, GA, USA (2000)
25. Rosemann, M., Recker, J.C., Flender, C.: Contextualisation of business processes. International Journal of Business Process Integration and Management 3 (1), 47–60 (2008)
26. Rosemann, M., Recker, J., Flender, C.: Designing context-aware Business Processes. In: Siau, K., Chiang, R., Hardgrave, B.C. (eds.) Systems analysis and design. People, processes and projects, pp. 51–73. M.E. Sharpe, Armonk, NY u.a (2011)
27. Soffer, P.: On the Notion of Flexibility in Business Processes. In: Proceedings of the CAiSE’05 Workshops, pp. 35–42 (2005)
28. Rosemann, M., van der Aalst, W.: A configurable reference modelling language. Information Systems 32, 1–23 (2007)
29. La Vara, J.L. de, Ali, R., Dalpiaz, F., Sánchez, J., Giorgini, P.: Business Processes Contextualisation via Context Analysis. In: Parsons, J., Saeki, M., Shoval, P., Woo, C., Wand, Y. (eds.) Conceptual modeling - ER 2010. 29th International Conference on Conceptual Modeling, Vancouver, BC, Canada, November 1 - 4, 2010 ; proceedings, 6412, pp. 471–476. Springer, Berlin (2010)
30. Saidani, O., Nurcan, S.: Towards Context Aware Business Process Modelling. In: Workshop on Business Process Modelling, Development, and Support, p. 1. Norway (2007)
31. Al-alshuhai, A., Siewe, F.: An Extension of Class Diagram to Model the Structure of Context-Aware Systems. In: The Sixth International Joint Conference on Advances in Engineering and Technology (AET) (2015)
32. Al-alshuhai, A., Siewe, F.: An Extension of UML Activity Diagram to Model the Behaviour of Context-Aware Systems. In: Computer and Information Technology; Ubiquitous Computing and Communications; Dependable,

Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM), pp. 431–437 (2015)
33. Heinrich, B., Schön, D.: Automated Planning of Context-aware Process Models. University of Münster, Münster, Germany (2015)
34. Conforti, R., La Rosa, M., Fortino, G., ter Hofstede, A.H., Recker, J., Adams, M.: Real-time risk monitoring in business processes. A sensor-based approach. Journal of Systems and Software 86, 2939–2965 (2013)
35. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Q 28, 75–105 (2004)
36. Object Management Group (OMG): Meta Object Facility (MOF) Core Specification (2006)
37. Atkinson, C., Gerbig, R., Fritzsche, M.: Modeling Language Extension in the Enterprise Systems Domain. In: 2013 17th IEEE International Enterprise Distributed Object Computing Conference (EDOC 2013), pp. 49–58
38. Seel, C.: Reverse Method Engineering. Methode und Softwareunterstützung zur Konstruktion und Adaption semiformaler Informationsmodellierungstechniken. Logos-Verl., Berlin (2010)
39. Karlsson, F., Ågerfalk, P.J.: Method configuration. Adapting to situational characteristics while creating reusable assets. Information and Software Technology 46, 619–633 (2004)
40. Mertens, P., Back, A. (eds.): Lexikon der Wirtschaftsinformatik. Springer, Berlin (2001)
41. Backus, J.W., Wegstein, J.H., van Wijngaarden, A., Woodger, M., Bauer, F.L., Green, J., Katz, C., McCarthy, J., Perlis, A.J., Rutishauser, H., et al.: Report on the algorithmic language ALGOL 60. Commun. ACM 3, 299–314 (1960)
42. Niesen, T., Dadashnia, S., Fettke, P., Loos, P.: A Vector Space Approach to Process Model Matching using Insights from Natural Language Processing. In: Nissen, V., Stelzer, D., Straßburger, S., Fischer, D. (eds.) Multikonferenz Wirtschaftsinformatik (MKWI) 2016. Technische Universität Ilmenau, 09. - 11. März 2016, vol. Ivol. , pp. 93–104. Universitätsverlag Ilmenau, Ilmenau (2016)
43. Weske, M.: Business process management. Concepts, languages, architectures. Springer, Berlin u.a. (2007)

Systematische Identifikation von Fachkomponenten mit SOM

Sebastian Schlauderer, Martin Robel, and Thomas Friedrich

University of Bamberg, Chair of Industrial Information Systems, Bamberg, Germany
{sebastian.schlauderer, martin.robel, thomas.friedrich}@uni-bamberg.de

Abstract. Aufgrund der stetig wachsenden Herausforderungen ist es für Unternehmen heutzutage von großer Bedeutung, Informationssysteme flexibel und kostengünstig anpassen zu können. Hierfür ist es jedoch unabdingbar, dass die Software-Bausteine eines Systems entsprechend systematisch abgeleitet und gestaltet werden. Die Business-Component-Identification-Methode (BCI-Methode) verspricht vor dem Hintergrund der Komponentenorientierung die Identifikation geeigneter Software-Bausteine zu gewährleisten. Ein Vorteil dieser Methode ist es, dass sie unabhängig von der Art der zugrundeliegenden konzeptionellen Modelle angewandt werden kann. Da die Qualität der Modelle jedoch die Güte der gefundenen Lösung maßgeblich determiniert, ist es indes fraglich, welche Modellarten sich besonders gut eignen. Die Methodik des Semantischen Objektmodells (SOM) bietet einen ganzheitlichen und etablierten Ansatz zur systematischen Beschreibung des Fachkonzepts. Im Rahmen des Artikels wird daher untersucht, inwiefern die Kombination von BCI- und SOM-Methodik zu validen Ergebnissen führt. Hierfür wird insbesondere untersucht, welche Modelle betrachtet werden müssen und welche Beziehungen zu unterscheiden sind.

Keywords: Komponentenorientierung, SOM-Methodik, BCI-Methode

1 Motivation

Die Entwicklung von Informationssystemen steht heutzutage vor einer Vielzahl teils konterkarierender Herausforderungen. Einerseits ist es notwendig, auf sich ändernde Markbedingungen schnell reagieren zu können und flexibel Funktionalität hinzufügen zu können [1]. Andererseits wächst der Kostendruck auf die Entwicklungsabteilungen [1]. Vor diesem Hintergrund verspricht die modulare Entwicklung von Informationssystemen eine Vielzahl an Vorteilen [2], [3]. Durch die Strukturierung von Informationssystemen in einzelne, lose gekoppelte Bausteine wird es ermöglicht, Informationssysteme flexibel anzupassen oder zu erweitern. Durch die Wiederverwendung dieser Bausteine, auch Fachkomponenten genannt [4], können weiterhin Entwicklungszeiten und Entwicklungskosten drastisch gekürzt werden [5]. Im Vergleich zur reinen Objektorientierung kann durch das Zusammenfassen mehrerer Objekte zu einer Fachkompo-

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Schlauderer, S.; Robel, M.; Friedrich, T. (2017): Systematische Identifikation von Fachkomponenten mit SOM, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 316-330

nente das Problem fragiler Basisklassen eingegrenzt werden und gleichermaßen sichergestellt werden, dass Vererbungsstrukturen das Geheimnisprinzip nicht durchbrechen [6].

Aufgrund der erwarteten Vorteile wurde eine Vielzahl an Methoden und Werkzeugen sowohl in der wissenschaftlichen Literatur als auch in der Praxis vorgeschlagen, welche die Konzeption, die Implementierung sowie die Assemblierung von Fachkomponenten zum Ziel haben [2], [6-9]. Insbesondere die Frage, wie ein Informationssystem möglichst effizient in Fachkomponenten partitioniert werden kann, ist von zentraler Bedeutung, da hierdurch die Anpassbarkeit des Systems sowie die Wiederverwendbarkeit der Fachkomponenten determiniert wird [10]. Einerseits kann die Wiederverwendbarkeit der Fachkomponenten erhöht werden, indem möglichst feingranulare Dienste zusammengefasst werden, die eine sehr spezifische Funktionalität anbieten [6]. Andererseits wird hierdurch jedoch die Anzahl an Fachkomponenten erhöht, was deren Assemblierung in einem Informationssystem erschwert [6].

Um diesen konfliktären Zielen gerecht zu werden, wurde unter anderem die sogenannte Business-Component-Identification-Methode (BCI-Methode) vorgeschlagen, die auf Basis von konzeptionellen Modellen eine heuristische Optimierung der Problemstellung vorsieht [11]. Ein Vorteil der Methode verspricht insbesondere zu sein, dass sie unabhängig von der Art der zugrundeliegenden Modelle verwendet werden kann. Die als Ausgangsbasis verwendeten konzeptionellen Modelle determinieren damit jedoch auch maßgeblich die Qualität der gefundenen Lösung. Dies wirft die zentrale Frage auf, welche Modelle und Notationen sich für die Komponentenfindung besonders gut eignen. Mit Blick auf die Etablierung komplexer, innovativer Geschäftslösungen als soziotechnische Systeme im Rahmen des Business Engineering [12] wächst die Bedeutung einer ganzheitlichen Betrachtung der Modellierung auf der strategischen, konzeptionellen sowie technischen Ebene. Das Semantische Objektmodell (SOM) stellt einen solchen ganzheitlichen Modellierungsansatz dar [13-16]. Die Unternehmensarchitektur des SOM-Ansatzes differenziert die drei Modellebenen Unternehmensplan, Geschäftsprozessmodell und Ressourcenmodell, welche jeweils aus einer struktur- sowie verhaltensorientierten Sichtweise in Form textueller Beschreibungen bzw. semi-formal durch Diagramme beschrieben werden [16]. Während das Vorgehensmodell die Schritte zur Erstellung der einzelnen Artefakte festlegt, stellen Metamodelle die Integration der Sichten sicher. Die konsequente Verfeinerung des Unternehmensplans hin zum Ressourcenmodell in einem Top-Down-Ansatz ermöglicht die Erstellung eines hochwertigen und vor allem hinsichtlich der Anforderungen vollständigen Fachkonzepts, das als Ausgangsbasis wesentlich zum Erfolg der Komponentenfindung beiträgt. Während der SOM-Ansatz also eine etablierte Methodik bereitstellt um das Fachkonzept detailliert zu beschreiben, unterstützt die BCI-Methode die Transformation von Fach- zu Systemkonzept. Trotz dieser vermeintlich symbiotischen Beziehung wurde das Zusammenspiel beider Methoden bisher nicht wissenschaftlich untersucht, weswegen sich insbesondere folgende Forschungsfragen ergeben:

1. Inwiefern eignet sich die SOM-Methodik zur Unterstützung der Modularisierung von Informationssystemen mithilfe der BCI-Methode?
2. Welche SOM-Modelle tragen konkret zur Erlangung valider Ergebnisse bei?

3. Welche Beziehungsformen lassen sich in den SOM-Modellen differenzieren und wie lassen sich diese bei entsprechender Eignung auf die BCI-Methode abbilden?

Zur Beantwortung der Forschungsfragen wird in Abschnitt 2 zunächst der theoretische Grundstock gelegt. Im Rahmen von Abschnitt 3 wird ein Fallbeispiel eingeführt, das die Anwendung der BCI-Methode auf Basis der verschiedenen SOM-Modelle demonstriert. Die Ergebnisse dieser prototypischen Anwendung werden anschließend in Abschnitt 4 diskutiert. Abschließend wird in Abschnitt 5 der Artikel kritisch zusammengefasst und ein Ausblick für weitere Untersuchungen gegeben.

2 Theoretische Grundlagen

In den folgenden Teilabschnitten werden grundlegende Begriffe komponentenbasierter Systeme, der BCI-Methode sowie der Modelle des SOM-Ansatzes erläutert.

2.1 Komponentenbasierte Systeme

Ein generelles Modell komponentenbasierter Anwendungssysteme wird im Common Business Component Model (COBCOM) erläutert [17]. Neben der Architektur komponentenbasierter Systeme wird hier zudem der Lebenszyklus von Fachkomponenten skizziert. Im Allgemeinen basieren komponentenbasierte Informationssysteme auf einer dreistufigen Client/Server-Architektur. Neben zusätzlichen technischen Anpassungen, die etwa der Vermittlung zwischen Dienstgebern und Dienstnehmern dienen, bspw. mithilfe sogenannter Object Request Broker, ist der zentrale Unterschied die weitere Differenzierung der Verarbeitungslogik [17]. Hierfür wird die betriebliche Funktionalität weiter in sogenannte Fachkomponenten aufgespalten.

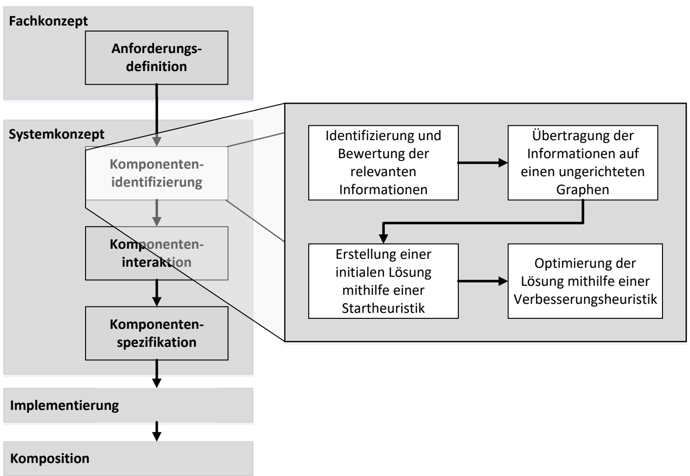


Abbildung 1. Entwicklungsprozess komponentenbasierter Systeme (links) [8] und Ablaufschritte der BCI-Methode (rechts)

Definition 1 [Fachkomponente]: „Eine Fachkomponente ist eine Komponente, die eine bestimmte Menge von Diensten einer betrieblichen Anwendungsdomäne anbietet" [4]. Fraglich ist jedoch, welche Menge an betrieblichen Diensten in einer Fachkomponente zusammengefasst wird. Der Entwicklungsprozess komponentenbasierter Systeme sieht daher auch einen eigenen Schritt zur Identifikation von Fachkomponenten vor (Abbildung 1). Es gilt den Identifikationsprozess so zu gestalten, dass die Anzahl an Schnittstellen zwischen den Fachkomponenten minimiert wird, während die Kohäsion innerhalb der Komponenten maximiert wird [6], [18]. Da die beiden Ziele jedoch konfliktär sind ist die Lösung so zu gestalten, dass beide berücksichtigt werden. Hierfür soll zunächst ein grundsätzliches Verständnis für die beiden Ziele geschaffen werden.
Definition 2 [Kopplung und Kohäsion]: Kopplung beschreibt das Ausmaß, in dem die Elemente innerhalb einer Fachkomponente mit Elementen außerhalb zusammenhängen. Kohäsion beschreibt das Ausmaß, in dem die Elemente innerhalb einer Fachkomponente in Wechselbeziehung zueinanderstehen. Die Kohäsion kann dabei als Kopplung innerhalb einer Komponente verstanden und gemessen werden [19], [20].

2.2 Methoden zur Identifikation von Fachkomponenten

Bei der Partitionierung eines Informationssystems und der Aufteilung der Dienste auf einzelne Fachkomponenten handelt es sich um ein NP-vollständiges Problem [21]. Zur Lösung dieses Problems wurden in der Literatur diverse Methoden vorgeschlagen und diskutiert [11], [22]. Viele dieser Ansätze schlagen allerdings nur grobe Richtlinen vor [23-27] und sind daher sehr stark abhänging von der Erfahrung des Anwenders. Im Vergleich dazu gibt die BCI-Methode dem Anwender zwar Freiheiten in der Gewichtung der verschiedenen Beziehungen, gibt jedoch auch ein systemgestütztes und einheitliches Vorgehen vor. Weiterhin gibt es diverse Methoden die auf Matrix Analysen basieren [28-30]. Während diese Ansätze zwar leicht anzuwenden sind, können mit ihnen nur zwei Dimensionen untersucht werden, wohingegen die BCI-Methode grundsätzlich beliebig viele Dimensionen berücksichtigt. Abschließend gibt es auch Ansätze, für die kein Tool-Support identifiert werden konnte, so dass diese Ansätze im Rahmen des Beitrages nicht anzuwenden waren [31], [32]. Folglich wurde auf die BCI-Methodik zurückgegriffen. Die Methode geht im Wesentlichen in vier Schritten vor (siehe Abbildung 1): Im ersten Schritt muss der Systemarchitekt entscheiden, welche Informationen aus den konzeptionellen Modellen gewonnen werden können und diese bzgl. ihrer Relevanz bewerten. Typischerweise handelt es sich bei den konzeptionellen Modellen um Geschäftsprozessmodelle, Datenmodelle oder um Funktionsdekompositionsmodelle. In diesem Schritt ist es anschließend notwendig, die verschiedenen Beziehungsarten zu unterscheiden und bezüglich ihrer Relevanz zu bewerten. Hiermit ist dem Systemarchitekten die Möglichkeit gegeben festzulegen, welche Beziehungsarten möglichst innerhalb der Fachkomponenten verlaufen sollen. Es handelt sich dabei um den zentralen Schritt, mit welchem der Systemarchitekt das spätere Ergebnis beeinflussen kann. Im zweiten Schritt werden alle als relevant identifizierten Informationen auf einen ungerichteten Graphen übertragen, wobei alle Elemente als Knoten und alle Beziehungen zwischen ihnen als Kanten abgebildet werden. Im dritten Schritt wird zunächst eine Startlösung generiert. Hierfür wird bspw. mithilfe der „Start-Partition-Greedy"-

Heuristik [11] eine Liste erstellt, welche alle Beziehungen bzgl. ihrer Relevanz sortiert. Am Ende dieses Schrittes steht eine initiale Aufteilung von Elementen zu Fachkomponenten. Im vierten Schritt soll diese Aufteilung optimiert werden, d.h. es soll die Kopplung minimiert und die Kohäsion maximiert werden. Hierfür ist es zunächst notwendig, die beiden Maßzahlen mathematisch zu definieren (Formel 1 und 2):

$$Cou(p) = \frac{1}{2}\sum_{i=1}^{k}\sum_{u\in C_i}\sum_{v\notin C_i} w_{(u,v)} \tag{1}$$

$$Coh(p) = \frac{1}{2}\sum_{i=1}^{k}\sum_{u\in C_i}\sum_{v\in C_i, v\neq u} w_{(u,v)} \tag{2}$$

Die Anzahl an Komponenten, die in der initialen Lösung gefunden wurden, wird durch die Variable *k* beschrieben. $w_{(u,v)}$ steht für das Gewicht, welches die Relevanz zwischen den beiden Elementen u und v ausdrückt. Im Rahmen des „Kernighan-Lin"-Algorithmus wird anschließend eine Optimierung der gefundenen Startlösung vorgenommen [33]. Einerseits wird die Summe der Gewichte der Beziehungen minimiert, die zwischen verschiedenen Fachkomponenten verlaufen (Formel 3 - links). Andererseits wird die Summe der Gewichte der Beziehungen maximiert, die innerhalb der Fachkomponenten verlaufen (Formel 3 - rechts). Unter der Nebenbedingung, dass die in der Startlösung gefundene Anzahl an Komponenten konstant bleibt, lässt es sich mathematisch nachweisen, dass die beiden Ziele mathematisch äquivalent sind [11].

$$min_p\ \frac{1}{2}\sum_{i=1}^{k}\sum_{u\in C_i}\sum_{v\notin C_i} w_{(u,v)} \iff max_p\ \frac{1}{2}\sum_{i=1}^{k}\sum_{u\in C_i}\sum_{v\in C_i, v\neq u} w_{(u,v)} \tag{3}$$

2.3 Modelle des SOM-Ansatzes

Das Semantische Objektmodell (SOM) stellt einen objekt- und geschäftsprozessorientierten Ansatz zur Modellierung betrieblicher Systeme dar und unterscheidet bis auf die Ebene des Unternehmensplans jeweils unterschiedliche Arten von Diagrammen zur Beschreibung von Struktur und Verhalten [15], [16]. Auf Ebene der Geschäftsprozesse bildet die Verknüpfung von betrieblichen Objekten durch Transaktionen im Interaktionsschema (IAS) die Struktur, die Verbindung von Aufgaben (Vorgangstypen) über Ereignisbeziehungen im Vorgangs-Ereignis-Schema (VES) das Verhalten ab. Analog beschreibt das konzeptuelle Objektschema (KOS) auf Ebene des Ressourcenmodells die Struktur des zu spezifizierenden Anwendungssystems durch konzeptuelle Objekttypen (KOT) sowie deren Interaktions-, Generalisierungs- und Aggregationsbeziehungen. Interaktionsbeziehungen zwischen Vorgangsobjekttypen (VOT) spezifizieren schließlich das Verhalten im Vorgangsobjektschema (VOS). Durch Integration der verschiedenen Modelle über die Sichten und Ebenen hinweg lässt sich sicherstellen, dass keine relevanten Informationen bei der Modellierung verloren gehen. Vielmehr findet eine Verfeinerung von einer abstrakten hin zu einer spezielleren Betrachtung auf Anwendugssystemebene statt. Das Vorgehensmodell sichert mit der Vorgabe von Zerlegungsregeln sowie der Integration der Sichten über ein Metamodell explizit die Konsistenz, Vollständigkeit und in der Folge die für die Anwendung der BCI-Methodik entscheidende Qualität des Fachkonzepts.

3 Komponentenfindung auf Basis des SOM-Ansatzes

Aufbauend auf den theoretischen Grundlagen wird nachfolgend die BCI-Methode auf Basis eines praxisnahen Fallbeispiels in SOM dargestellt. Der Fokus liegt dabei auf der Identifizierung und Bewertung relevanter Informationen aus den SOM-Modellen.

3.1 Einführung Fallsbeispiel OnlineCars

Das folgende Fallbeispiel stellt vereinfacht die Vermietung und Auftragsabwicklung von Elektrofahrzeugen über ein Internetportal (OnlineCars) dar. Mithilfe der SOM-Methodik wurden schrittweise die Schemata der einzelnen Ebenen erstellt.

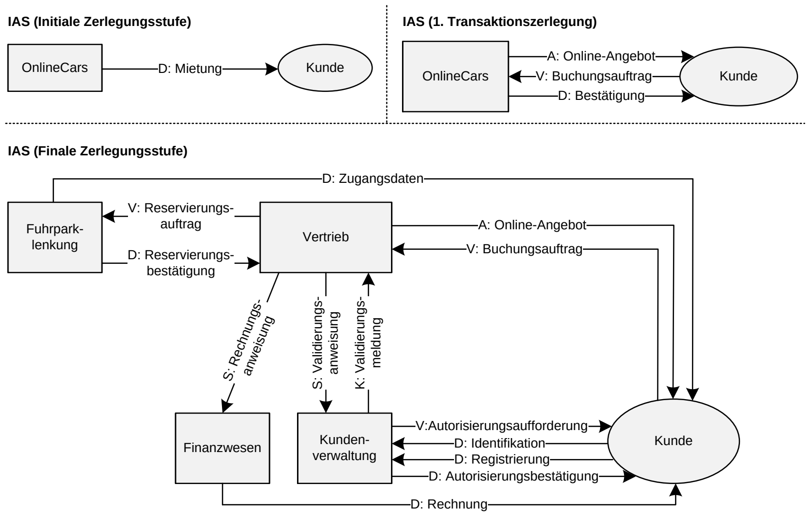


Abbildung 2. IAS des Fallbeispiels

Auf Geschäftprozessebene wird der Verleihservice in Abbildung 2 strukturorientiert in Form mehrerer Interaktionsschemata (IAS) dargestellt. Ausgehend vom initialen IAS wurden sukzessive Objekt- und Transaktionszerlegungen gemäß den Zerlegungsregeln des SOM-Ansatzes durchgeführt [16]. Während der Vertrieb dem Finanzwesen und der Kundenverwaltung hierarchisch übergeordnet ist, erfolgt die Koordination zwischen Vertrieb und Fuhrparklenkung sowie zwischen Vertrieb und dem Umweltobjekt Kunde auf Basis des Verhandlungsprinzips. Der Verleihservice erstellt als Leistung bei Vorliegen eines Buchungsauftrags neben der Autorisierungsbestätigung und der Rechnung auch die Zugangsdaten und übergibt diese dem Kunden. Die Leistungen werden nach Beauftragung bzw. Anweisung des Vertriebs von unterschiedlichen Objekten erbracht. Der Kunde übermittelt als Leistung aufgrund einer Autorisierungsaufforderung seine Registrierungs- bzw. Identifikationsdaten. In Ergänzung zum IAS beschreibt der Ausschnitt des VES in Abbildung 3 das Verhalten der betrieblichen Objekte im Geschäftsprozess. Der Empfang eines Buchungsauftrags durch den Vertrieb löst bspw. über ein objektinternes Ereignis die Sendung eines Validierungsauftrags an die Kundenverwaltung aus.

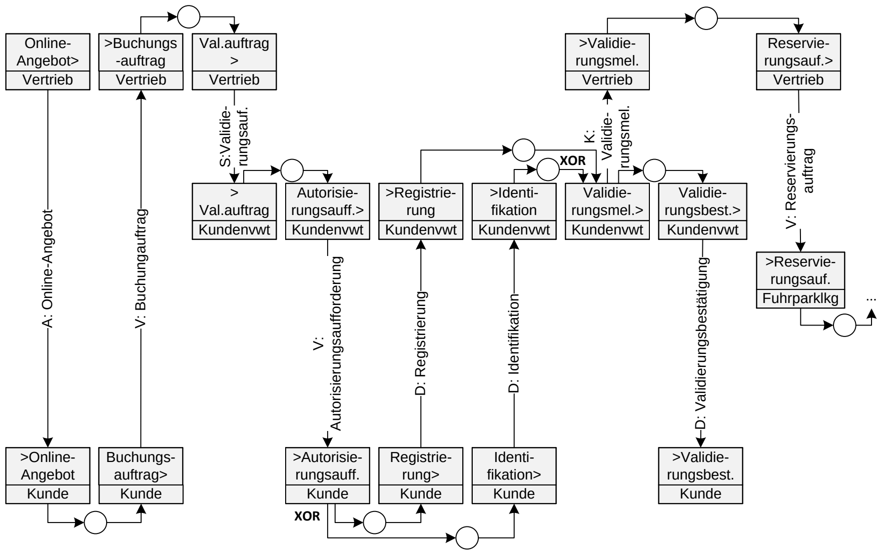


Abbildung 3. Ausschnitt VES des Fallbeispiels (finale Zerlegungsstufe)

Die Ableitung der Anwendungssystemspezifikation auf Basis der Geschäftsprozessmodelle IAS/VES in Form des KOS aus Struktursicht und des VOS aus Verhaltenssicht schließt die Modellierung des Anwendungsfalls ab. Abbildung 4 stellt das initiale und das konsolidierte KOS dar. Entsprechend des SOM-Ansatzes werden bei der Konsolidierung diejenigen KOTs entfernt, die mit nicht-automatisierten Aufgaben und Transaktionen korrespondieren [16]. Zudem werden KOTs, die sich in ihren Attributen und/oder Operatoren weitestgehend decken aus Redundanzgründen zusammengefasst. Im Fallbeispiel entstehen dabei die KOTs Autorisierung (*Validierungsauftrag*, *Autorisierungsaufforderung*) und Reservierung (*Reserverierungsauftrag* und *-bestätigung*). Es wurden keine KOTs entfernt, da sämtliche Aufgaben bzw. Transaktionen von dem spezifizierten Anwendungssystem unterstützt werden sollen. Zusätzlich werden im konsolidierten KOS die Beziehungsarten Interaktion (interacts_with), Aggregation (is_part_of) und Generalisierung (is_a) berücksichtigt.

Das für jedes betriebliche Objekt erstellte VOS wird in einer integrierten Darstellung in Abbildung 5 zusammengeführt. Analog zum KOS werden im VOS diejenigen VOTs entfernt, deren zugehörige Aufgaben nicht automatisiert werden. Ebenso werden VOTs zusammengefasst, deren Aufgaben aus Gründen der semantischen Integrität stets gemeinsam durchzuführen sind. Aus Gründen der Übersichtlichkeit wurden für das KOS und VOS des Fallbeispiels keine weitere Spezifikation, wie bspw. Zuordnung von Operatoren oder Definition von Nachrichten, vorgenommen. Die erzeugten Modelle auf der Geschäftsprozessebene (IAS/VES) sowie auf der Ebene der Anwendungssystemspezifikation (KOS/VOS) bilden die Grundlage für die Identifikation der relevanten Informationen im nächsten Abschnitt.

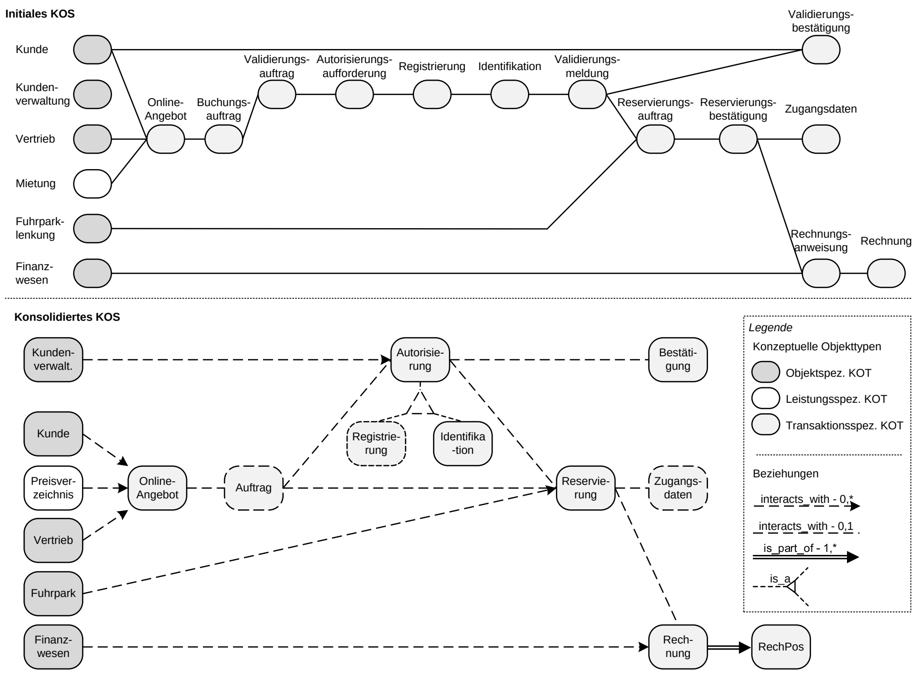


Abbildung 4. Initiales und konsolidiertes KOS des Fallbeispiels

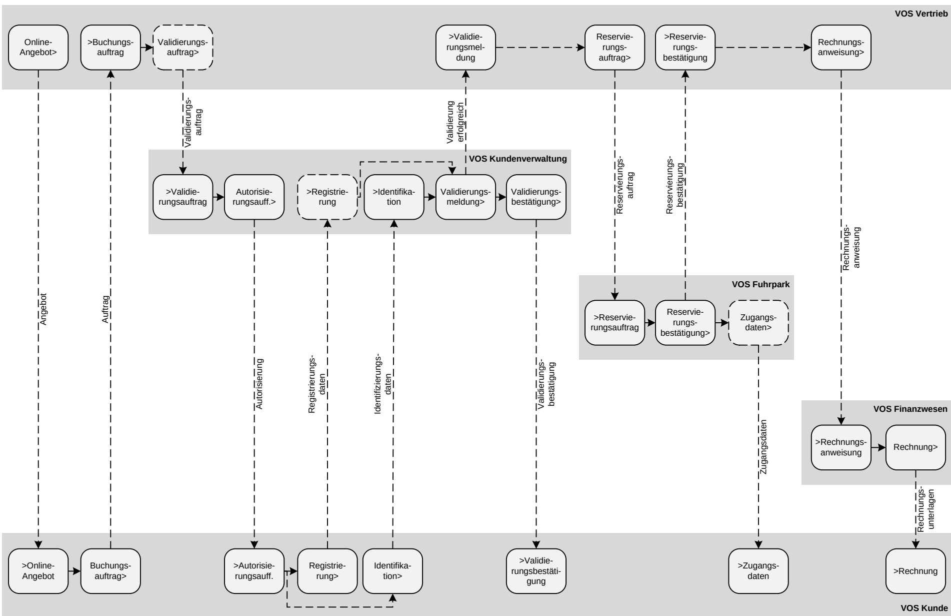


Abbildung 5. Konsolidiertes VOS des Fallbeispiels

3.2 Identifizierung und Bewertung der relevanten Informationen

Die Ergebnisse der Modellierung des Anwendungsfalls verdeutlichen bereits die Herausforderung bei der Entscheidung, welche Informationen bei der BCI-Methode zu berücksichtigen sind. Zum einen beziehen sich die Modelle des SOM-Ansatzes auf unterschiedliche Ebenen. Zum anderen wird innerhalb einer Ebene zwischen struktur- und verhaltensorientierten Modellen unterschieden. Zudem bauen die Modelle konsequent aufeinander auf und beschreiben damit ähnliche Sachverhalte.

Im Unterschied zu IAS und VES beinhalten KOS und VOS ausschließlich automatisierbare Elemente, wodurch Komponenten hervorgebracht werden können, die sich auch tatsächlich realisieren lassen. Gleichwohl entsteht aufgrund der Ableitung von KOS und VOS aus IAS und VES kein Informationsverlust, vielmehr wird eine Verfeinerung und Ergänzung hin zur Anwendungsspezifikation vorgenommen [16]. KOS und VOS bilden somit das Fachkonzept als Ausgangsbasis für die Komponentenfindung. In der systematischen Ausrichtung auf das Anwendungssystem als automatisierter Teil des betrieblichen Systems kann somit eine der Stärken von SOM ausgenutzt werden.

In einem weiteren Schritt werden die in KOS/VOS enthaltenen Beziehungen im Hinblick auf ihre Relevanz bewertet. Aufgrund der engen Kopplung von KOTs werden die Aggregations- und Generalisierungsbeziehungen im KOS relevanter eingestuft als Interaktionsbeziehungen zwischen KOTs. Zumindest theoretisch ließen sich zudem objektspezifische, leistungsspezifische und transaktionsspezifische KOTs differenzieren. Allerdings lässt sich nur schwer eine logische Begründung für die Einstufung der Relevanz geben. Interaktionsbeziehungen zwischen KOTs werden nicht berücksichtigt. Die Interaktionsbeziehungen im VOS lassen sich in Anlehnung an das VES nach objektinternen (objektinterne Ereignisse) und objektexternen Beziehungen (Transaktionen) differenzieren. Objektinterne Beziehungen verbinden Aufgaben bzw. VOTs innerhalb eines bestimmten betrieblichen Objekts und werden für die Komponentenfindung daher relevanter angesehen als objektexterne Beziehungen, die Aufgaben bzw. VOTs unterschiedlicher betrieblicher Objekte verbinden. Optionale Beziehungen (XOR), wie sie bspw. zwischen den VOTs >*Autorisierungsaufforderung* und *Registrierung*> bzw. *Identifikation*> bestehen, wird die geringste Relevanz zugewiesen. Objektexterne Beziehungen können als Teil einer Transaktion einer bestimmten Phase (Anbahnung, Vereinbarung, Durchführung) zugeordnet werden. Die Bewertung der Relevanz der verschiedenen Phasen lässt sich auch hier schwer logisch begründen, sodass objektexterne Beziehungen ohne weitere Differenzierung aufgenommen werden.

3.3 Übertragung der identifizierten Informationen auf einen Graphen

Die relevanten Informationen aus KOS und VOS müssen vor der Anwendung von Heuristiken der BCI-Methode auf einen Graphen übertragen werden. KOTs und VOTs werden als Elemente in Form von Knoten im Graphen abgebildet. Die Beziehungen zwischen den KOTs und zwischen den VOTs werden als ungerichtete Kanten dargestellt. Die fehlende Angabe der expliziten Zuordnung zwischen KOTs und VOTs in einem der Modelle führt zunächst zur Entstehung zweier unverbundener Teilgraphen. Es ist daher zur Erzeugung eines zusammenhängenden Graphen zu klären, welche KOTs mit welchen VOTs in Beziehung stehen. Grundsätzlich ist jeder VOT einem betrieblichen Objekt zugeordnet, dessen Name mit dem Namen des zugehörigen VOS korrespondiert. Die betrieblichen Objekte korrespondieren wiederum mit den entsprechenden objektspezifischen KOTs im KOS. Zunächst werden daher die einzelnen VOTs mit den objektspezifischen KOTs verbunden. Im Fallbeispiel korrespondiert bspw. der VOT *Online-Angebot>* mit dem VOS *Vertrieb*. Das VOS *Vertrieb* korrespondiert wiederum mit dem betrieblichen Objekt Vertrieb und dem objektspezifischen KOT *Vertrieb*. Es lässt sich damit eine Beziehung zwischen dem VOT *Online-Angebot>* und dem KOT *Vertrieb* herstellen. Aufgrund der im KOS durchgeführten Konsolidierung besteht jedoch die Möglichkeit, dass KOTs umbenannt oder mit anderen KOTs zusammengefasst wurden. Auf Grundlage der vorgenommenen Konsolidierungsschritte kann so bspw. der VOT *Rechnungsanweisung>* dem KOT *Rechnung*, welcher durch die Zusammenfassung der KOTs *Rechnungsanweisung* und *Rechnung* entstanden ist, zugeordnet werden. Nach der Verbindung der VOTs mit den KOTS veranschaulicht der in Abbildung 6 dargestellte Graph das Ergebnis der Übertragung der Objekte und Beziehungen.

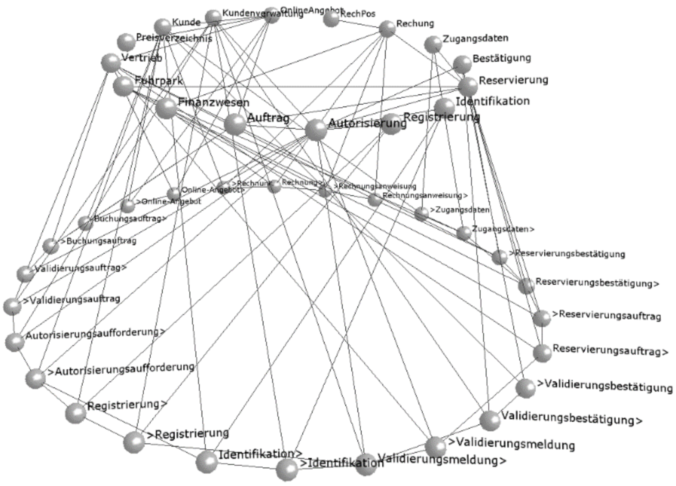


Abbildung 6. Graph der BCI-Methode

4 Ergebnisse des Fallbeispiels

Das Ergebnis der Komponentenfindung ist in Abbildung 7 als optimale Partitionierung des Graphen hinsichtlich der angewandten Heuristiken dargestellt. Die Aufteilung der KOTs und VOTs auf die verschiedenen Komponenten wurde durch die Anwendung der Start-Partition-Greedy-Startheuristik in Verbindung mit der Kernighan-Lin-Verbesserungsheuristik erreicht. Grundlage der Anwendung der Heuristiken bildet die Gewichtung der Beziehungen im Graphen. Dabei besteht die Herausforderung darin, die in Abschnitt 3.2 beschriebene unterschiedliche Relevanz der jeweiligen Beziehungen als konkrete numerische Gewichte abzubilden. Analog zur Festlegung der Relevanz der Beziehungen bestehen auch hier Freiheitsgrade für den Systemarchitekten. Um eine möglichst optimale Lösung zu generieren, wurden verschiedene Gewichtungen unter Beachtung der Relevanz der Beziehungen mithilfe eines Software-Tools (BCI-3D) getestet und die Ergebnisse hinsichtlich ihrer Eignung diskutiert. Die dem Ergebnis zugrundeliegenden Gewichtungen können Abbildung 7 entnommen werden.

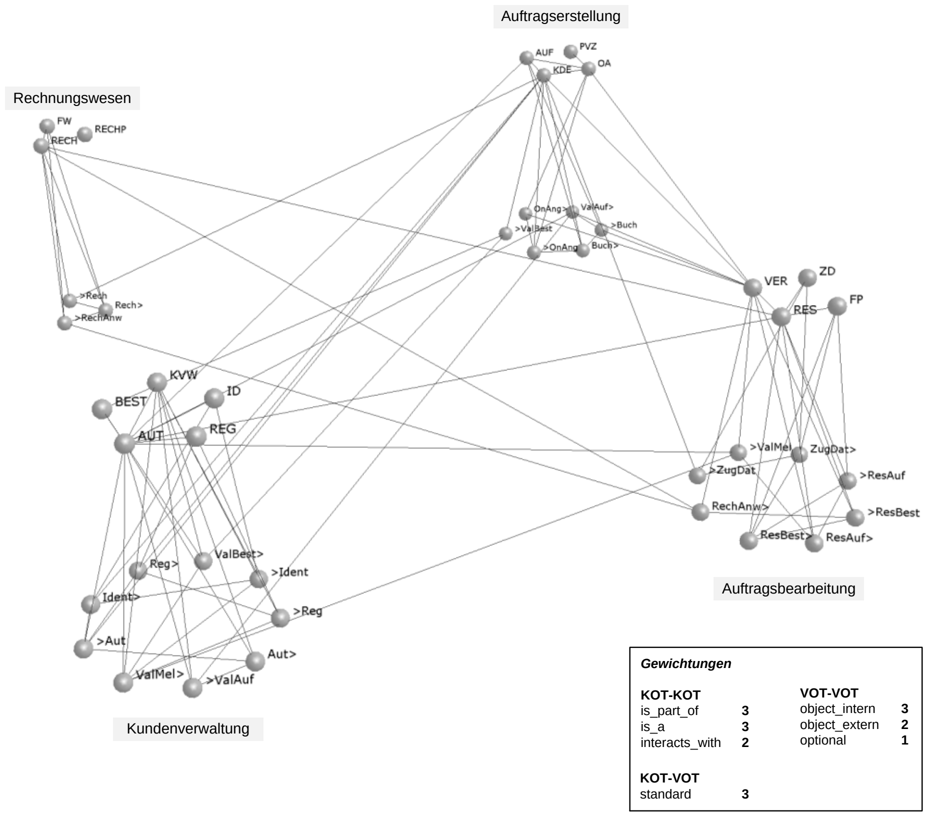


Abbildung 7. Optimierte Partitionierung des Graphen

Sofern keine Abstufungen aufgrund einer Rangfolge vorgesehen sind, wurde ein Standardgewicht von 3 gewählt. Folglich erhalten bspw. die den objektinternen Beziehungen untergeordneten objektexternen Beziehungen eine geringere Gewichtung von 2. Dem schließen sich die optionalen Beziehungen mit einem Gewicht von 1 an. Insgesamt wurden mit der vorgeschlagenen Gewichtung vier verschiedene Komponenten identifiziert. Jede Komponente fasst im Sinne der Erzielung der maximalen Kohäsion zusammenhängende KOTs und VOTs zusammen und wird in Bezug auf die Gemeinsamkeiten der bereitgestellten Dienste mit einem entsprechenden Namen versehen. Die Komponente *Auftragserstellung* veröffentlicht das Angebot an Fahrzeugen und erstellt die Buchungsaufträge der Kunden. Die Verwaltung der Kunden sowie deren Autorisierung durch Identifikation und Registrierung ist Aufgabe der Komponente *Kundenver-*

waltung. Die Reservierung von Fahrzeugen aus dem Fuhrpark, die Bestätigung der Reservierung sowie die Übermittlung der Zugangsdaten wird von der Komponente *Auftragsbearbeitung* angeboten. Schließlich stellt die Komponente *Rechnungswesen* die Funktionalität für die buchhalterische Erfassung des Auftrags und die Erstellung der Rechnungsdokumente bereit. Die verbleibenden Beziehungen zwischen den Komponenten realisieren in Form von Schnittstellen die Kopplung der Komponenten. Betrachtet man die Kosten der Partitionierung, so konnten diese nach Anwendung der Kernighan-Lin-Heuristik von $\mathrm{C(P)} = 101$ der Startlösung auf $\mathrm{C(P)} = 70$ reduziert werden. Konkret ergibt dies ein Kopplungsverhältnis von $\overline{\mathrm{Cou(P)}} \approx 0.20$ sowie entsprechend ein Kohäsionsverhältnis von $\overline{\mathrm{Coh(P)}} \approx 0.80$. Im Vergleich zur Startlösung konnte damit im Sinne einer Optimierung der Partitionierung das Kopplungsverhältnis um 0.09 verringert und das Kohäsionsverhältnis um 0.09 erhöht werden. Konkret bedeutet dies, dass die Zahl der Verbindungen zwischen den Komponenten im Vergleich zur Startlösung nochmals reduziert werden konnte.

5 Zusammenfassung und Ausblick

Im vorliegenden Beitrag wurde aufgezeigt, wie eine Komponentenfindung nach der BCI-Methode auf Basis des SOM-Ansatzes durchgeführt werden kann. Zunächst wurden hierfür die Grundlagen der Komponentenfindung, die Vorgehensweise der BCI-Methode sowie die Modelle des SOM-Ansatzes erläutert. Anschließend wurde die Anwendung der BCI-Methode auf Basis des SOM-Ansatzes anhand eines praxisnahen Fallbeispiels dargestellt. Die Ergebnisse des Fallbeispiels zeigen, wie ausgehend von den verschiedenen Modellen des SOM-Ansatzes ein Anwendungssystem in eine möglichst optimale Menge an Fachkomponenten unterteilt werden kann. In Bezug auf die in diesem Beitrag adressierten Forschungsfragen lassen sich folgende Rückschlüsse ziehen. Der objekt- und prozessorientierte SOM-Ansatz stellt mit seinen verschiedenen, aufeinander aufbauenden Modellen eine geeignete Ausgangsbasis für die Komponentenfindung dar. Für die Komponentenfindung eignen sich hierbei insbesondere die Modelle auf der Ebene der Anwendungssystemspezifikation, da diese die von einem Anwendungssystem zu unterstützenden Informationen beinhalten. Im Hinblick auf die Unterscheidung und Übertragung der für die Komponentenfindung relevanten Beziehungen bietet der SOM-Ansatz verschiedene Möglichkeiten. Für das Fallbespiel wurden bspw. die verschiedenen Beziehungsarten innerhalb des KOS als unterschiedlich relevant für die Komponentenfindung angesehen. Ebenso wurden die Interaktionsbeziehungen innerhalb des VOS weiter in objektinterne, objektexterne und optionale Beziehungen differenziert und in dieser Reihenfolge mit einer absteigenden Relevanz für die Komponentenfindung versehen. Durch die Trennung von strukturorientierten und verhaltensorientierten Modellsichten innerhalb des SOM-Ansatzes ergibt sich bei der Übertragung der relevanten Informationen auf einen Graphen die Problematik, dass hierbei zwei nicht miteinander verbundene Teilgraphen entstehen. Um dieser Problematik entgegenzuwirken, wurde innerhalb des Beitrags eine Möglichkeit aufgezeigt, wie beide Teilgraphen auf Grundlage der Informationen aus KOS und VOS zu einem gemeinsamen Graphen zusammengeführt werden können. Durch das Erstellen einer initialen Lösung mittels einer Startheuristik und der Optimierung dieser Lösung mittels einer Verbesserungsheuristik konnte eine möglichst optimale Aufteilung von vier Fachkomponenten ermittelt werden. Jede Fachkomponente wird dabei von einer Menge aus KOTs und VOTs repräsentiert, die gegenüber den KOTs und VOTs der anderen Fachkomponenten eine maximale Kohäsion und eine minimale Kopplung aufweisen.

Aus den Ergebnissen und Limitationen des Beitrags ergeben sich sowohl für die Forschung als auch für die Praxis mehrere Implikationen. Innerhalb des Beitrags wurde ein einzelnes, vereinfachtes Fallbeispiel betrachtet, bei welchem aus Gründen der Einfachheit und Übersichtlichkeit keine weiterführende Spezifikation der SOM-Modelle auf der Ebene der Anwendungsspezifikation vorgenommen wurde. Aus Sicht der Praxis bezieht sich der Nutzen vor allem auf die Möglichkeit, systematisch Komponenten auf Grundlage eines objektorientierten Ansatzes wie SOM identifizieren zu können. Mithilfe des im Beitrag vorgestellten Vorgehens wird erstmals eine methodische Grobstruktur bereitgestellt, die erläutert in welchen Schritten SOM-Modelle zur Komponentenfindung mit der BCI-Methode eingesetzt werden können. Weiterhin kann in zukünftigen praxisnahen Forschungsarbeiten die Unterscheidung der Beziehungen sowie die Gewichtungen der Beziehungsarten diskutiert werden, um eine konsolidierte Unterscheidung der Beziehungen zu erreichen. Auf Basis eines solchen konsolidierten Ergebnisses kann ein Systemarchitekt weitaus leichter Anpassungen entsprechend seiner persönlichen Präferenzen vornehmen oder auch die Sensitivität der Ergebnisse prüfen. Zudem kann mithilfe des verwendeten Tools, das die BCI-Methode umsetzt, der Ansatz auch in größeren und industrienahen Projekten verwendet werden. Insbesondere bei solchen Projekten verspricht die vorgestellte Methode mithilfe des Tools eine Vielzahl an Vorteilen gegenüber einer manuellen Systemaufteilung mit sich zu bringen.

Aus Sicht der Forschung könnte bspw. die Übertragung des Vorgehens auf weitere, komplexere Fallbeispiele einen Mehrwert dahingehend bieten, die Eignung der im Beitrag ausgewählten Beziehungen inkl. deren Relevanz zu überprüfen. In diesem Zusammenhang könnte es ebenfalls von Interesse sein, weitere Differenzierungen der Beziehungsarten zu diskutieren und diese hinsichtlich ihrer Praktikabilität zu untersuchen. Potenzial besteht im Besonderen mit Blick auf die Zusammenführung von KOS und VOS bei der Berücksichtigung der Zuordnung von Operatoren zu einzelnen VOTs sowie bei der Definition von Nachrichtenbeziehungen als Erweiterung der Anwendungsspezifikation in SOM. Grundsätzlich ist die Integration der Modellierungssichten über ein Metamodell eine der Stärken des SOM-Ansatzes. Im Zusammenspiel mit der Vorgabe definierter Zerlegungsregeln in Form eines Vorgehenesmodells kann die Konsistenz sowie die Vollständigkeit des Fachkonzepts sichergestellt werden. Die systematische Ausrichtung auf das Anwendungssystem als automatisierter Teil ist weiterhin kongruent zu den Zielen der BCI-Methode. Inwiefern Anwender hierdurch zu hochwertigeren Fachkonzepten gelangen, und damit im Rahmen der BCI-Methode bessere Ergebnissen bzgl. Kohäsion und Kopplung erzielen, müsste jedoch in zukünftigen Forschungsarbeiten empirisch geprüft werden. Weiterhin können als Resultat solcher Diskussionen SOM-Modelle auch in anderen Komponentenfindungsverfahren verwendet werden, die ebenfalls zwischen den diskutierten Beziehungen unterscheiden müssen.

Ein Merkmal zur Reduktion der Abhängigkeiten bei der Komponentenorientierung bildet die Erhöhung des Abstraktionsniveaus bei der Entwicklung, da Komponenten in

Bezug zum Klassenkonzept der Objektorientierung mehr Funktionalität kapseln und dadurch abstraktere Entwurfseinheiten darstellen [34], [35]. Hinsichtlich der Gemeinsamkeit zwischen Klassen und Komponenten durch die Kapselung von Funktionalität konnte im vorliegenden Beitrag aufgezeigt werden, wie sich systematisch mithilfe der BCI-Methode Elemente des objektorientierten SOM-Ansatzes zu semantisch sinnvollen, abstrakteren Einheiten in Form von Komponenten zusammenfassen lassen. Die Ergebnisse des SOM-Ansatzes stellen dabei mit der bereits vorhandenen Abgrenzung einzelner Objektklassen eine sehr gute Ausgangsbasis für den Ausbau der Anwendungsspezifikation hin zu einem komponentenbasierten System dar. Während die SOM-Modelle einen bewährten Ansatz liefern, um das Fachkonzept detailliert zu beschreiben, kann darauf aufbauend die BCI-Methode verwendet werden, um das Systemkonzept abzuleiten. Die BCI-Methode liefert eine systematische Unterstützung für den Übergang vom Fach- zum Systemkonzept durch die Identifikation geeigneter Fachkomponenten, ist hierbei jedoch auf hochwertige Modelle im Fachkonzept angewiesen. Insofern konnte im vorliegenden Beitrag gezeigt werden, dass sich die beiden Ansätze ideal ergänzen und weiteren Boden für zukünftige Forschungsarbeiten liefern.

Literatur

1. McDonald, M., Begin, J., Fortino, S.: Meeting the challenge: the 2009 CIO agenda. Gartner, Inc. (2009)
2. Herzum, P., Sims, O.: Business Component Factory: A Comprehensive Overview of Component-Based Development for the Enterprise. John Wiley & Sons, NY (2000)
3. Vitharana, P., Zahedi, F., Jain, H.: Knowledge-Based Repository Scheme for Storing and Retrieving Business Components: A Theoretical Design and an Empirical Analysis. IEEE Transactions on Software Engineering 29, 649-664 (2003)
4. Turowski, K.: Fachkomponenten: Komponentenbasierte betriebliche Anwendungssysteme. Shaker, Aachen (2003)
5. Sharp, J., Ryan, S.: A Review of Component-Based Software Development. In: 26th International Conference on Information Systems (ICIS) (2005)
6. Szyperski, C., Gruntz, D., Murer, S.: Component Software: Beyond Object-Oriented Programming. Addison-Wesley, Harlow (2002)
7. Brown, A.W.: Large-Scale, Component-Based Development. Prentice Hall, Upper Saddle River, NJ (2000)
8. Cheesman, J., Daniels, J.: UML Components: A Simple Process for Specifying Component-Based Software. Addison-Wesley, Upper Saddle River, NJ (2001)
9. Lim, W.C.: Managing Software Reuse - A Comprehensive Guide to Strategically Reengineering the Organization for Reusable Components. Prentice Hall, Upper Saddle River, NJ (1998)
10. Gorton, I.: Essential Software Architecture. Springer, Heidelberg (2011)
11. Birkmeier, D., Overhage, S.: A Method to Support a Reflective Derivation of Business Components from Conceptual Models. ISeB 11, 403-435 (2012)
12. Österle, H., Blessing, D.: Ansätze des Business Engineering. HMD - Praxis der Wirtschaftsinformatik 241, 7-17 (2005)
13. Ferstl, O.K., Sinz, E.J.: Objektmodellierung betrieblicher Informationssysteme im Semantischen Objektmodell. WIRTSCHAFTSINFORMATIK 32, 566-581 (1990)

14. Ferstl, O.K., Sinz, E.J.: Ein Vorgehensmodell zur Objektmodellierung betrieblicher Informationssysteme im Semantischen Objektmodell (SOM). BISE 33, 477-491 (1991)
15. Ferstl, O.K., Sinz, E.J.: Der Ansatz des Semantischen Objektmodells (SOM) zur Modellierung von Geschäftsprozessen. BISE 37, 209-220 (1995)
16. Ferstl, O.K., Sinz, E.J.: Grundlagen der Wirtschaftsinformatik. Oldenbourg, München (2013)
17. Rautenstrauch, C., Turowski, K.: Common Business Component Model (COBCOM): Generelles Modell komponentenbasierter Anwendungssysteme. In: 5. Internationale Tagung Wirtschaftsinformatik, pp. 681-695. (2001)
18. Parnas, D.L.: On the Criteria to be Used in Decomposing Systems into Modules. Communications of the ACM 15, 1053-1058 (1972)
19. Vitharana, P., Jain, H., Zahedi, F.: Strategy-Based Design of Reusable Business Components. IEEE Trans Syst Man Cybern 34, 460-474 (2004)
20. Wallnau, K.C., Hissam, S.A., Seacord, R.C.: Building Systems from Commercial Components. Addison-Wesley, Upper Saddle River, NJ (2002)
21. Garey, M., Johnson, D., Stockmeyer, L.: Some Simplified NP-Complete Problems. In: 6th Annual ACM Symposium on Theory of Computing, pp. 47–63 (1974)
22. Birkmeier, D., Overhage, S.: On Component Identification Approaches – Classification, State of the Art, and Comparison. In: Lewis, G.A., Poernomo, I., Hofmeister, C. (eds.) Component-Based Software Engineering, vol. 5582, pp. 1-18. Springer, Berlin, Germany (2009)
23. Sugumaran, V., Tanniru, M., Storey, V.: Identifying software components from process requirements using domain model and object libraries. In: 20th International Conference on Information Systems (1999)
24. Levi, K., Arsanjani, A.: A goal-driven approach to enterprise component identification and specification. Communications of the ACM 45, 45-52 (2002)
25. Cui, J., Chae, H.: Applying agglomerative hierarchical clustering algorithms to component identification for legacy systems. Information and Software Technology 43, 601-614 (2011)
26. Kang, K., Kim, S., Lee, J., Kim, K., Shin, E., Huh, M.: FORM: a feature-oriented reuse method with domain-specific reference architectures. Ann Softw Eng 5, 143-168 (1998)
27. Birkmeier, D.Q., Gehlert, A., Overhage, S., Schlauderer, S.: Alignment of Business and IT Architectures in the German Federal Government: A Systematic Method to Identify Services from Business Processes. In: Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS), pp. 3848-3857 (2013)
28. Kim, S., Chang, S.: A systematic method to identify software components. In: 11th Asia-Pacific software engineering conference (2004)
29. Lee, S., Yang, Y., Cho, E., Kim, S., Rhew, S.: COMO: A UML-based component development methodology. In: 6th Asia-Pacific software engineering conference (1999)
30. Jang, Y., Kim, E., Lee, K.: Object-oriented component identification method using the affinity analysis technique. In: 9th Conference OOIS (2003)
31. Ganesan, R., Sengupta, S.: A technique for the design of component-based applications. In: 39th Conference TOOLS (2001)
32. Meng, F., Zhan, D., Xu, X.: Business component identification of enterprise information system: a hierarchical clustering method. In: IEEE international conference on e-Business engineering (2005)
33. Kernighan, B., Lin, S.: An Efficient Heuristic Procedure for Partitioning Graphs. Bell System Technical Journal 49, 291-307 (1970)
34. Sommerville, I.: Software Engineering. Pearsons Education, Harlow (2007)
35. Zwintzscher, O.: Software-Komponenten im Überblick: Einführung, Klassifizierung & Vergleich von JavaBeans, EJB, COM+,. Net, CORBA, UML 2. W3L, Herdecke (2005)

Towards A Method for Developing Reference Enterprise Architectures

Felix Timm[1], Kurt Sandkuhl[1], Michael Fellmann[1]

[1]University of Rostock, Chair of Business Information Systems, Rostock, Germany
{felix.timm,kurt.sandkuhl,michael.fellmann}@uni-rostock.de

Abstract. In most economic sectors organizations face rapid environmental changes like regulations. Such changes can force them to adjust both their organizational and operational structure. For instance, in the energy utility sector numerous developments moved German Public Utilities (PUs) towards a liberalized market. Nowadays PUs have to stay competitive while managing a heterogeneous information technology (IT) landscape. We address this demand for aligning business and IT by combining the holistic perspective of Enterprise Architecture Management (EAM) with the characteristic of reference modeling to reuse knowledge in a problem domain. Therefore, we utilize configurative reference modeling within Design Science Research (DSR). The artefact at hand is a method for developing a Reference Enterprise Architecture (R-EA), which is applied in the problem domain of PUs. Our contributions are the (i) adaptation of Configurative Reference Modelling (CRM) to develop a R-EA and (ii) a procedure how to elicit knowledge for R-EA development method.

Keywords: Reference Modelling; Public Utility Industry; Enterprise Architecture Management; Reference Enterprise Architecture

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Timm, F.; Sandkuhl, K.; Fellmann, M. (2017): Towards A Method for Developing Reference Enterprise Architectures, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 331-345

1 Introduction

After the directives by the Council of the European Union from 1996 [1] and 2003 [2], Europe's energy utility industry has faced significant structural market changes in terms of market liberalization and renewable energy sources [3]. A PU is defined a natural or legal entity that supplies energy, operates an energy grid or holds power if disposition on these grids [4]. In Germany, PUs suddenly faced many challenges after the government passed the German Energy Act [4]. They do not only imply organizational or operational changes for PUs but also business opportunities for third parties to intensify competition in a prior monopolistic market. Nowadays PUs face the need to integrate new technologies resulting in heterogeneous IT landscapes. For instance, flexible demand-side management by the integration of Smart Grid technologies becomes necessary due to the volatility and uncertainty of renewable energy resources like wind or solar energy [5]. Moreover, the installation of new smart and real-time metering systems requires the reorganization of PUs' metering data management [6, 7]. This requires adaption of the PU's business models and the reengineering of existing and the establishment of new business processes [3]. Furthermore, integrating new technologies like smart metering systems has an impact on the PUs' IT architecture [7]. A survey with 53 respondents from German PUs' management level revealed that organizations currently lack in fulfilling this requirement of effective Business-IT alignment, although they assess it vital to cope with the challenged stated above [8].

In order to tackle these different challenges sketched above, it is not sufficient to focus on single aspects or initiatives. Rather, a holistic approach is needed such as EAM that creates a holistic perspective on the organization and is capable of revealing relationships between the strategy of an organization, business processes, responsible roles, applications and information infrastructures [9]. Implementing an EAM initiative would allow PUs to locate environmental changes within their organization in order to reveal necessary Business-IT alignment actions. Despite this potential, recent information systems research (ISR) activities hardly addresses EAM aspects, since it primarily focuses on Smart Grid technologies or the market itself [10], [11, 12]. They do not create holistic models or methods tailored to the PUs' situation in particular. Still, some show the multi-dimensional consequences and opportunities that challenges like smart metering cause for PUs [13]. Although practitioners agree on the potential of EAM discussed above, they consider implementing such initiatives from scratch as too resource-consuming [8]. To address this issue, we suggest the construction of a reference model as a useful approach to support PUs in benefiting from the EAM discipline. Such a model would enable organizations to reuse knowledge and reduce the effort of creating an individual model, e.g. a concrete model of an enterprise architecture [14].

In order to support the construction of such a model, we devise a method to develop a Reference Enterprise Architecture (R-EA) that is based on established literature found in the ISR sub-disciplines of reference modeling and EAM. In more detail, we adapt a configurational reference modeling method [15] in the frame of a Design Science Research approach [16]. The resulting artefact, the R-EA method, is focused on data elicitation as well as on representational aspects. The work is structured as follows:

After discussing reference modeling approaches and related work in section 2, our application of configurative reference modeling within DSR is presented in section 3. In section 4 we present the R-EA method. The method's evaluation and application in a Use Case is presented in sections 5 and 6 before concluding our work on section 7.

2 Background

As depicted earlier, practitioners agree with us that EAM would support PUs to overcome the stated market challenges [8]. This is in line with current research, since it is agreed that establishing EAM within organizations through architectural thinking increases manageability and flexibility as well as consistency [17]. EAM is a management discipline to understand, plan, develop, control and adjust organizational structures. Therefore different perspectives are developed resulting in a holistic view, which captures the current state as well as a target state of the organization [9]. These perspectives are manifold and a plethora of frameworks exist how to structure and develop such an EAM initiative [18]. In our work we use the TOGAF framework [19] as it is widely accepted among practitioners and comes with an detailed modeling language specification ArchiMate [20] as well as an open source modeling tool Archi[1].

In our work, we leverage the principles of the TOGAF framework in our method for reference model construction, the R-EA method. This means that concrete reference models constructed using our method should conform to the principles and structure of the TOGAF framework. In order to provide more background connected to the development of the R-EA method which is at the core of the paper at hand, we discuss general reference modeling approaches in the next Section 2.1. After this, we discus related work and define our work's problem statement in Section 2.2.

2.1 Utilizing Reference Modeling Approaches

A reference model can be defined as "every model […] which can be used in supporting the construction of another model can be seen in this sense as a reference model" [14], p.491). This notion emphasizes a use-oriented perspective. However, in order to be reusable, other authors such as [21] argue that also other characteristics such as the universality and recommendation have to be considered while developing a reference model. This is in line with our aim of developing a reference model that is intended for reuse. With regard to reuse, [15] identify a dilemma in reference modelling among the general validity of the model during construction and the effort of adjusting the model during its application. Their approach suggests to solve this conflict by developing *configurative reference models*, which define rules to determine model adjustments during application. This approach integrates the application aspects into the construction phase of reference modelling. We adopt this technique since the configuration of the presented R-EA would allow a certain PU to tailor the R-EA to its individual structure and thus avoiding to develop a model from scratch or adapting an

[1] See http://www.archimatetool.com.

extensive model that may be a resource consuming and tedious task. In the main part, this configurational aspect is discussed in more detail.

Regarding the overall approach of reference modelling, older works see reference model construction and application as an integrated process (e.g. [22, 23]). This is in contrast with more recent works differentiating between a phase of model *construction* and model *application* (e.g. [14]). We stick to the former distinction and present a method for the construction of a R-EA. This implies that the application of the models that can be created with this method is part of our future work.

In order to develop a reference modelling method, it is important to distinguish between two generic strategies of reference model development. While the deductive developed models emerge by deriving from generally accepted knowledge, the inductive approach abstracts from individual models to agree on a common understanding within the reference model [24]. As a part of our work on the R-EA, we propose an initial approach how to combine these two strategies into a hybrid reference modeling development method [25]. This is made transparent in Section 4.1.

Regarding these considerations, we understand the R-EA method developed in the research underlying the paper at hand as a method which supports the development of reference models that ease the implementation of EAM initiatives at German PUs. The models created using the R-EA method should be reusable, universal and should give recommendations towards the PUs on how to structure their organizations depending on certain market roles.

2.2 Related Work and Problem Statement

Only few research activities can be identified that relates to our purpose of a R-EA development method. As a starting point to develop a R-EA in the financial industry, [26] propose a working definition on enterprise reference architectures corresponding to the authors' term R-EA. They describe an R-EA as "… a generic EA for a class of enterprises, that is a coherent whole of EA design principles, methods and models which are used as foundation in the design and realization of the concrete EA that consists of three coherent partial architectures: the business architecture, the application architecture and the technology architecture" [26]. Further, [27] developed a reference requirements set in the e-government domain, which serves as a starting point for certain EAM initiatives in this domain. The approach utilizes the motivation extension of TOGAF and ArchiMate and aims to develop a complete R-EA for the e-government domain. Despite these works do identify the need for a method, neither of them addresses methodical aspects.

Most literature in the utility domain address the context of environmental sustainability, but lacks in investigating the implications for PUs' IS and its role in the current developments [28]. There has been little research activity directly addressing EAM as a means to face the challenges in the utility industry. Most research focuses on parts of EAM's scope only. [29] identified 11 reference models for information systems development in utility industry and proposed a catalogue for reference models in order to agree on a common terminology. Other topics addressed are Smart Grids [10, 11], Smart Metering Systems [6], load management and demand response [30] or dynamic tariff models [31]. All these research activities address issues PUs nowadays have to consider not only in their business but also in their information systems. The stated literature investigates this at a relatively detailed level.

We point out that current literature lacks a methodical approach for developing R-EAs. Other R-EA related research in domains such as public administration or the financial industry do not clarify how to elicit necessary knowledge for a R-EA in general nor how a R-EA should be structured. Based on these considerations this work addresses the following problem statement.

Problem Statement. We identify the need for a R-EA development method that guides data elicitation and representation of the model. The problem is addressed by two separate research questions (RQs):

RQ1: How should a R-EA method support the elicitation of the necessary domain knowledge for developing the R-EA?

RQ2: How could a R-EA method support the model structure development and the R-EA model evaluation?

3 Methodological Considerations

We selected the Design Science Research (DSR) paradigm as our research design since it supports the development of artefacts as a contribution to tackle real-world problems. At the same time, it facilitates deriving conclusions relevant for research [16]. This decision is also taken by other authors who developed reference models in diverse areas such as Master Data Management [32] or Supply Chain Management [33]. This is no surprise since DSR is essentially a problem-solving paradigm that seeks to improve the reality, in which enterprises act, by creating and analyzing innovative IT artefacts such as models, methods and prototypes [16]. Consequently, the DSR paradigm is highly relevant for method development. Our work refers to the DSR approach proposed by [34]. In this model, a regulative cycle is defined comprising the phases of *(I) problem investigation, (II) solution design, (III) design validation* and *(IV) solution implementation*, which are traversed iteratively. We applied the DSR regulative cycle as presented in Figure 1. In phase (I) we conducted an online survey, that revealed the need for a R-EA method in the certain problem domain of PUs [8]. We further analyzed literature towards methodical approaches of R-EA development and derived our problem statement as discussed in sections 1 and 2. Within (II) we adapted the method for configurative reference modeling [15] to the concepts of EAM. Therefore, we extended the CRM method with specifics of TOGAF. This resulted in an additional procedure for data collection and guidelines for structuring a developed R-EA, which are explained in section 4. In phase (III) the method was validated using Technical Action Research [35] by applying the R-EA method in the PU problem domain described above. This is illustrated in section 6. The phase (IV) is out of the scope of this paper and will be subject of our future work.

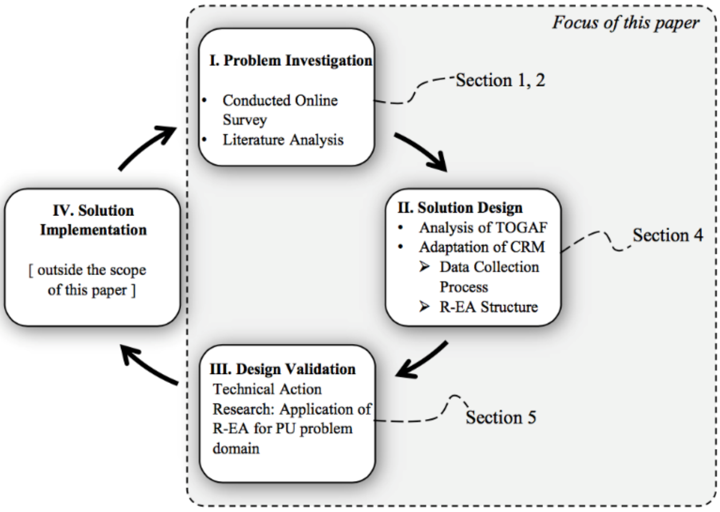


Figure 1. Applied DSR Regulative Cycle according to [34]

4 Method for R-EA development

We developed the R-EA development method based on the CRM method proposed by [36] and adapted towards the specifics of EAM. CRM comprises the five phases for reference modeling. In (1) the project's objective is defined. Therefore, the scope of the R-EA has to be clarified by identifying the problem domain and the class of organizations that are addressed by the R-EA. Also requirements towards the R-EA are set. An example of the R-EA problem domain could be the integration of smart metering systems at PUs. During phase (2) the modeling approach needs to be defined. For R-EA development one need to define what EAM framework (e.g. TOGAF) and modeling language (e.g. ArchiMate) to use. Further, the configurational aspects of the R-EA are defined. These two phases set up phase (III) were the actual reference modeling is conducted. The data is collected and the R-EA is modeled. The R-EA is refined by evaluating it in phase (IV). In the end, the R-EA is released to the market. Figure 2 illustrates that our R-EA method extends CRM in phases (3) and (4). We assess the CRM to be not comprehensive enough regarding data collection and model structure for R-EA model development. Phase (5) is out of the scope of this paper.

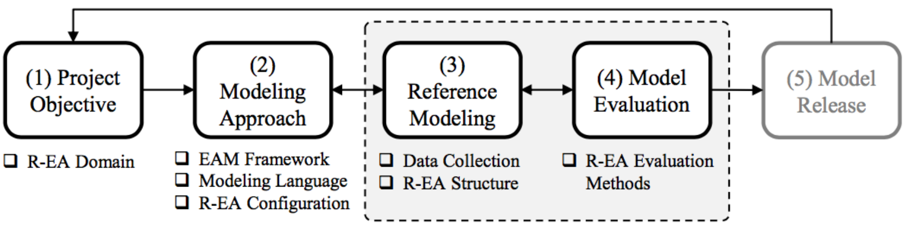


Figure 2. Basic R-EA development method applying CRM [36]

4.1 R-EA Method: A Procedure for Data Collection during R-EA Development

As described above we extend the reference modeling task CRM's phase (3) in order to comply with specifics of EAM. Therefore, we add a procedure model to phase (3), which proposes how to elicit the relevant data in order to perform R-EA development later in this phase. The procedure comprises three steps shown in Figure 3. Each step uses several data collection methods and produces an output in the form of a R-EA model version. Section 4.2 explains how to define the R-EA's elements and structure is defined.

Our R-EA method uses a hybrid approach of deductive (Step A) and inductive (Step B) reference modeling before abstract the data to final R-EA (Step C). In *Step (A)* an initial R-EA is developed using deductive collection methods. In order to ensure objective findings we propose to use methodological triangulation according to [37]. For instance, the data sources could include an analysis of domain-specific literature, expert interviews with practitioners of the domain as well as a survey addressing representatives of the prior defined organization class. The analysis of this data produces an **initial R-EA**, following a structural guideline described in section 4.2.

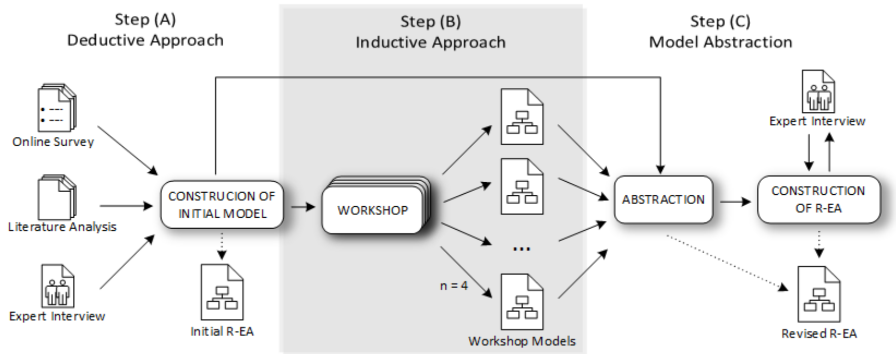


Figure 3. Development Process of the R-EA

Step (B) defines to conduct inductive reference modeling by collecting concrete enterprise architectures of representative organizations from the problem domain. The organizations' current state and challenges are captured with focus on the defined R-EA problem domain. To do so the R-EA method suggests to conduct workshops at the representative organizations'. A concrete workshop design is proposed in [38]. The workshop, its preparation and post-processing is in line with the first three stages of preparation, collection and preprocessing for inductive reference modeling from [39]. The workshop has to collect data from different perspectives on the R-EA's problem domain. These perspectives depend on the EAM framework that was chosen in (II) of CRM. When using TOGAF, elements from business, data, application and technology layer have to be considered [19]. Therefore, we recommend to conduct several small workshops at each organization. They differ in their focus and participants from the particular organization. In order to enable effective data collection, each workshops should be conducted following the guidelines of participative enterprise modeling [40].

There are several roles participating in a workshop. The research team consists of the modeling facilitator and minute takers. While the modeling facilitator moderates, asks questions and illustrates the talking points, minute takers keep record of the discussion. The participants from the organization, i.e. the domain experts, depend strongly on the focus of the respective workshop. For instance, identifying the business areas of a PU that are influenced by the integration of smart metering systems requires domain experts to be from the management level. In contrast, collecting information regarding concrete business processes or data models for smart meter installation rather requires process owners or specialists as participants. Thus, the workshops have to be planned wisely and a concrete agenda needs to be communicated to the organization. After condensing the collected data and cross-checking it with the participants, an individual **workshop model** is developed for each organization using ArchiMate. These have to be comparable and, thus, must comply with the modeling conventions defined by the researchers. This is discussed in section 4.2.

In *Step (C)*, the actual R-EA is developed based on the initial R-EA and the workshop models. It is then validated by dint of expert interviews resulting in a **revised R-EA**. For abstracting to the R-EA, reference modeling research identifies clustering as a means to derive a reference model, which requires the comparability of the individual models [39]. This is also applicable towards the R-EA. Still, research only addresses the comparability of process models in this regard and lacks methodical aspects for comparing individual EA models. In this last step of model abstraction, it is vital that the initial and all workshop models are comparable with each other, since they are more complex than process models. Therefore, the next section explicates how to agree on a structure for the R-EA development.

4.2 R-EA Method: Definition of R-EA Structure

As depicted above it is essential to model comparable individual EA models. Therefore, we suggest how to define a concrete R-EA structure that guides the development of each model (i.e. initial R-EA, workshop model and revised R-EA). Therefore, it is important to understand how an EA model is structured. First, a meta model defines what elements are used to model the EA model, e.g. the TOGAF meta-model [19]. Second, different types of relationships are defined how these elements are related among each other, e.g. a certain role *is assigned to* a business process. Third, viewpoints are used as projections on parts of the EA model to visualize a certain purpose of the model, e.g. to see what business processes are using a certain software component [20]. We understand a sound R-EA structure to consider all of these three aspects in order to guarantee comparability of the individual models.

The definition of a R-EA structure is influenced by *(a) the problem domain* and intention of the R-EA identified during (I) of CRM, *(b) the meta model* of the chosen EAM framework as well as its modeling language and *(c) the data available* for conducting the R-EA method. We suggest the following actions to follow in order to develop a R-EA structure. Keeping the problem domain of the R-EA method application in mind, the meta model of the choses EAM framework has to be analyzed. It has to be clarified what EA layers to cover, what concrete EA elements as well as what relationship types to use. This results in a modified meta model. Afterwards, a list of viewpoints has to be defined that explicates what elements and relationships are projected in what viewpoint for what purpose. When using the TOGAF framework, the ArchiMate specification provides a profound documentation of the meta-model and defines standard viewpoints [20]. Although the R-EA structure should be agreed on before modeling the initial R-EA model, data quality and data availability during data collection in steps (A) and (B) from Figure 3 will influence the R-EA structure. Thus, it may be modified during the R-EA development.

4.3 R-EA Method: Evaluation of the R-EA

Next to the construction perspective on R-EA development, our R-EA method provides a procedure how to evaluate the developed R-EA. The procedure is illustrated by Figure 4. Next to the data collection the workshops can be used to validate and refine the initial R-EA model. After conducting the model abstraction step, the resulting R-EA needs to be further validated. The R-EA method addresses this by conducting another expert interview. We recommend to interview specialists for evaluating detailed parts of the R-EA. In order to assess the R-EA's overall validity managers or business consultants should be interviewed. On the one hand, the interviewees should be familiar with EA models. On the other hand, they should not have been played a part in the R-EA construction.

The R-EA can be interpreted as a set of statements formulated in a modelling language. It can be evaluated first according to whether they are built properly (e.g. adhering to the syntactical and grammatical rules) and second to their content (e.g. semantical meaningfulness, truthful, understandable). Thus, we suggest the evaluation activities to distinguish between construction- and content-centric aspects. The construction-centric perspective concentrates on whether modeling conventions of the chosen EAM framework were followed and design decisions for the R-EA structure are appropriate towards the R-EA problem domain. Addressing the content-centric perspective it has to be assesses whether the R-EA is valid regarding semantic aspects (e.g. conciseness, consistency, completeness) and pragmatic aspects (e.g. appropriateness). During the workshops and interviews, the participants should be confronted with the model and judge it according to these aspects.

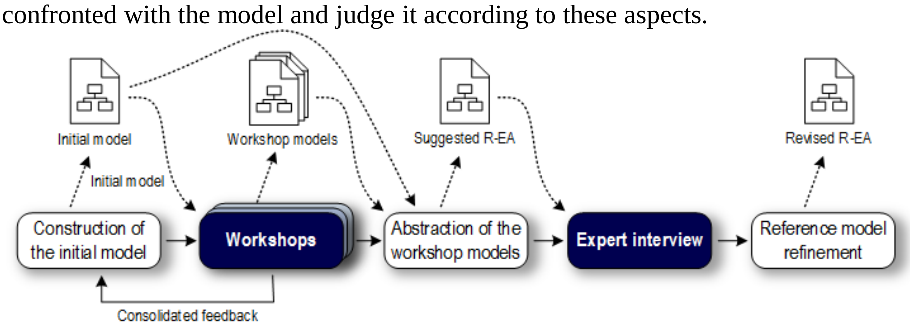


Figure 4. Procedure for Evaluating the R-EA

5 Validation of the R-EA Method

As visualized in Figure 1 we applied technical action research (TAR) for design validation of our R-EA development method [35]. After developing a first version of the method, we applied it to a problem domain. Before summarizing the method application in section 6, we discuss the setting of TAR application and want to make the R-EA method's evolution more transparent.

The R-EA method was applied in order to develop a R-EA that identifies the consequences of the market liberalization and the emerge of renewable energy sources for PU organizations. We worked together with a vendor of an enterprise resource planning (ERP) system in the energy sector, with whom we together conducted the workshops at four German PUs. During the R-EA method application (i.e. during the definition of the R-EA structure, after each workshop and during the modeling of the R-EA model) as well as after the method application we conducted interviews with the experts of the ERP system vendor addressing benefits and drawbacks of the R-EA method. According to [34] we were able to assess the method's internal validity and try to make first conclusions regarding its external validity. This relates to the observational evaluation method for DSR [16].

The method presented in section 4 represents the latest version of the R-EA method. During its application in the PU problem domain, necessary modifications of the method were identified and incorporated later on. Referring to Figure 3, the conduction of expert interviews was added in Step (A) in the first place. With the help of the vendor's expert knowledge of the industry, the initial R-EA model gain in quality and was then enriched with information from the other sources. Secondly, Step (B) was enhanced in terms of the workshop design. After we defined a strict workshop agenda, the experts noticed that the workshops need to be more flexible regarding the independence of the participants' availability. Third, in Step (C) we added the evaluation loop by conducting another expert interview in order to evaluate the final R-EA model towards construction- and content-centric aspects. Fourth, the importance of a sound definition of the R-EA structure became clear during the modeling activities in Step (B). This led to a more precise process how to structure the R-EA (see section 4.2).

According to [41] the evaluation of our artefact is primarily of formative nature. The application of the R-EA had a significant impact on the method's structure. A final expert interview with the project partner assessed the R-EA method as internal valid as defined in [34]. The developed R-EA met the partner's expectations and was used in subsequent meetings with their clients, i.e. PUs. In order to make a credible statement regarding external validity, the R-EA method needs to be applied in another problem domain. We currently apply it in a project located in the finance sector and first results indicate that the method is applicable to a different context.

6 Applying the R-EA Method to the Problem Domain of PUs

As described above the R-EA method was applied to the use case of the PUs current situation. The application was conducted together with a project partner, an ERP system vendor specialized on PUs. In Phase (I) (see Figure 1) of CRM the problem objective was to develop a R-EA that identifies the consequences of the market liberalization and the emerge of renewable energy sources for PU organizations. The class of addressed organization were PUs in Germany. The national restriction was due to the fact, that we only had positive feedback from PUs located in Germany. In (II) we agreed on using the TOGAF framework in line with the modeling language ArchiMate. The models were created with the open source tool Archi, which has full ArchiMate support. During studying the initial literature, it became clear that the business models of German PUs strongly depend on the market role they take. Thus, we derived that the R-EA will be configurable by the market role a certain PU takes. In (III) Step (A) we conducted an online survey [8], expert interviews with specialists from the partner, and analyzed domain-specific literature. This resulted in an initial R-EA, which was validated again with the partner's experts. In (III) Step (B) we designed the workshop agenda for collecting the data at four PUs. Each workshop was conducted on two days and comprised of five different workshops slots, using a top-down approach. In the first slot PU's participants were managers and department chiefs (i.e. IT department, corporate development). The focus was to identify the PU's organizational structure, its market roles as well as their business-to-customer and business-to-business relationships. For the remaining four slots, it was decided to pick four domain-specific business processes (namely customer acquisition, meter data collection, consumption-based billing and house connection) and discuss each of them with the respective process owners of the particular PU. While the first slot focused on TOGAF's business layer, the latter four collected more detailed information also for the data and application layer.

In parallel to (III) Step (A) and (B) we developed a R-EA structure and defined modeling conventions according to section 4.2. After the analysis of the data from Step (A) we modified the TOGAF meta-model by concentrating on the most vital EA elements like business actor, function, application component or data object. Due to the problem domain, we excluded all elements from the technology layer. On the basis of this modified meta-model we identified four viewpoints to model each EA model for the R-EA development: *business function, actor cooperation, service realization (for each of the four business processes) and application structure viewpoint* [20]. Three project members modeled the workshop models. Following the defined R-EA structure, it was possible to minimize the different individual modeling intentions, that occurred in the beginning. During project internal modeling workshops, we further agreed on name conventions for the same phenomena.

In (III) Step (C) the R-EA was derived by analyzing the initial R-EA and the four workshop models. During internal modeling workshops we identified similarities regarding element types, their names and relationships with other elements for each viewpoint. This was conducted manually and thus assessed as very time-intense. Figure 5 visualizes an extract of the final R-EA, the service realization viewpoint of *"contract closing"*. A PU provides this service to the consumer during the customer acquisition business process. The model shows which processes (e.g. customer registration), IT systems (e.g. CRM software), business objects (e.g. Core Data of the consumer), additional services (e.g. Checkout former Supplier) and business events (e.g. form arrives) are related to the realization of closing a new contract with a consumer. It thus also reveals the interplay between TOGAF's business, data and application layer.

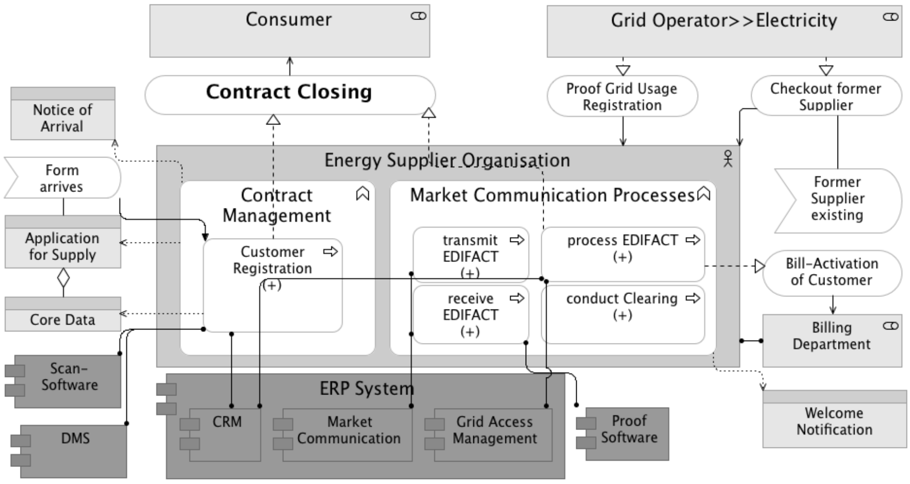


Figure 5. Extract of Service Realization Viewpoint for the Business Process Contract Closing

7 Discussion and Conclusion

Our work proposes a method for reference enterprise architecture (R-EA) development within a design science research (DSR) framework [16]. Therefore, the artefact builds on the configurative reference modeling (CRM) method [36] and extends it. Using technical actions research (TAR) [35], we apply the R-EA method in the problem domain of the energy industry and discuss the method's evolution. Our first contribution is the *(i) adaption of Configurative Reference Modelling to develop a R-EA*. We extend the CRM method and apply it for the R-EA development. Further, we suggest *(ii) a procedure how to elicit knowledge for R-EA development method*. These contributions answer RQ1. Thus, an approach for data collection is developed. We present a workshop design using participative modelling [40]. Next to knowledge elicitation the R-EA method extends CRM by providing an approach for defining a R-EA structure as well as an evaluation design of the resulting R-EA model (both addressing RQ2).

After one iteration of the regulative cycle [34], the R-EA method was assessed internally valid as the developed R-EA met the partner's expectations and was used in subsequent meetings with their clients, i.e. PUs. In terms of external validity resp. sensitivity analysis, we cannot make a credible statement so far, since the R-EA method needs to be applied to other problem domains. Currently, we apply it in a project located in the finance sector and first results indicate that the method is applicable to a different context.

Next to the validation results, there are several limitations of the R-EA method we want to point out. Although we stick to TAR by [35], the validation design holds a bias risk, since we as the method designers also applied it for validation. For future work, we should rather act as observers when applying the method in another use case. This leads to the limitation that we assess, that the method holds too much implicit knowledge for its application. This will be addressed by using a method conceptualization (e.g. [42]) and developing a method handbook. Further, the current version of the R-EA method mainly focuses eliciting the current state of practice in the problem domain, e.g. how PUs work at the moment in the changing market. Expanding this focus to future requirements of the problem domain would incorporate a *to-be* R-EA model. According to the experts this would raise the value of a resulting R-EA model. Moreover, the configuration aspect of the R-EA method needs to be investigated in more detail. This would enable to broaden the R-EA method's focus towards the application of a developed R-EA model a certain organization. Finally, also economic aspects of R-EA application have to be investigated.

References

1. (1996) DIRECTIVE 96/92/EC
2. (2003) DIRECTIVE 2003/54/EC
3. Appelrath H-J, Chamoni P (2007) Veränderungen in der Energiewirtschaft — Herausforderungen für die IT. Wirtsch. Inform. 49(5): 329–330. doi: 10.1007/s11576-007-0076-8
4. (2005) German Energy Act (Energiewirtschaftsgesetz): EnWG
5. Goebel C, Jacobsen H-A, del Razo V et al. (2014) Energy Informatics. Bus Inf Syst Eng 6(1): 25–31. doi: 10.1007/s12599-013-0304-2
6. Vukmirovic S, Erdeljan A, Kulic F et al. (2010) A smart metering architecture as a step towards Smart Grid realization. In: IEEE International Energy Conference (ENERGYCON 2010), pp 357–362
7. Aichele C, Dalkmann U, Margardt P et al. (2009) Business Process Framework and IT Architecture for Smart Meter Reading. In: Wirtschaftsinformatik Proceedings, vol 247, pp 647–656
8. Timm F, Wißotzki M, Köpp C (2015) Current State of Enterprise Architecture Management in SME Utilities. In: Cunningham DW (ed) Informatik 2015: Tagung vom 28. September – 02. Oktober 2015 in Cottbus. Ges. für Informatik, Bonn, pp 895–907
9. Ahlemann F, Stettiner E, Messerschmidt M et al. (2012) Strategic Enterprise Architecture Management. Springer Berlin Heidelberg, Berlin, Heidelberg
10. Appelrath H-J, Beenken P, Bischofs L et al. (2012) IT-Architekturentwicklung im Smart Grid: Perspektiven für eine sichere markt- und standardbasierte Integration erneuerbarer Energien. Springer Berlin Heidelberg, Berlin, Heidelberg
11. Smart Grid Coordination Group (2012) Smart Grid Reference Architecture: CEN, CENELEC,ETSI.
12. Beer S, Rüttinger H, Bischofs L et al. (2010) Towards a Reference Architecture for Regional Electricity MarketsEntwurf einer Referenzarchitektur für regionale Elektrizitätsmärkte. it - Information Technology 52(2). doi: 10.1524/itit.2010.0572

13. Jagstaidt UCC, Kossahl J, Kolbe LM (2011) Smart Metering Information Management. Bus Inf Syst Eng 3(5): 323–326. doi: 10.1007/s12599-011-0173-5
14. Thomas O (2006) Understanding the Term Reference Model in Information Systems Research: History, Literature Analysis and Explanation. In: Bussler C (ed) Business process management workshops: BPM 2005 international workshops, BPI, BPD, ENEI, BPRM, WSCOBPM, BPS, Nancy, France, September 5, 2005 ; revised selected papers, vol 3812. Springer, Berlin, pp 484–496
15. Becker J, Delfmann P, Dreiling A et al. (2004) Configurative Process Modeling - Outlining an Approach to Increased Business Process Model Usability. In: Proceedings of Information Resources Management Association Conference (IRMA) 2014, Orleans, New Orleans, pp 615–619
16. Hevner AR, March ST, Park J et al. (2004) Design science in information systems research. MIS Quarterly 28(1): 75–105
17. Winter R (2014) Architectural Thinking. Bus Inf Syst Eng 6(6): 361–364. doi: 10.1007/s12599-014-0352-2
18. Matthes D (2011) Enterprise Architecture Frameworks Kompendium. Springer Berlin Heidelberg, Berlin, Heidelberg
19. The Open Group (2010) TOGAF Version 9, 9. ed. [Nachdr.]. TOGAF series. Van Haren Publishing, Zaltbommel
20. The Open Group (2015) ArchiMate® 2.1 specification: Open Group Standard, Second edition, second impression, Mai 2015. The Open Group Series. Van Haren Publishing, Zaltbommel
21. Vom Brocke J (2003) Referenzmodellierung: Gestaltung und Verteilung von Konstruktionsprozessen. Univ., Diss.--Zugl.: Münster, 2002. Advances in information systems and management science, vol 4. Logos, Berlin
22. Schütte R (1998) Grundsätze ordnungsmäßiger Referenzmodellierung. Gabler Verlag, Wiesbaden
23. Fettke P, Loos P (2004) Referenzmodellierungsforschung. Wirtschaftsinf 46(5): 331–340. doi: 10.1007/BF03250947
24. Becker J, Schütte R (1997) Referenz-Informationsmodelle für den Handel: Begriff, Nutzen und Empfehlungen für die Gestaltung und unternehmensspezifische Adaption von Referenzmodellen. In: Krallmann H (ed) Wirtschaftsinformatik ’97: Internationale Geschäftstätigkeit auf der Basis flexibler Organisationsstrukturen und leistungsfähiger Informationssysteme. Physica-Verlag HD, Heidelberg, pp 427–448
25. Jana-Rebecca Rehse, Philip Hake, Peter Fettke et al. (2016) Inductive Reference Model Development: Recent Results and Current Challenges. In: Heinrich C. Mayr, Martin Pinzger (eds) INFORMATIK 2016. Jahrestagung der Gesellschaft für Informatik (INFORMATIK-2016), September 26-30, Klagenfurt, Austria, P-259. GI, Bonn
26. ten Harmsen van der Beek, Wijke, Trienekens J, Grefen P (2012) The Application of Enterprise Reference Architecture in the Financial Industry. In: Aier S, Ekstedt M, Matthes F et al. (eds): 7th Workshop, TEAR 2012, and 5th Working Conference , Barcelona, Spain, October 23-24, 2012. Proceedings, vol 131. Springer, Berlin, Heidelberg, pp 93–110
27. Tambouris E, Kaliva E, Liaros M et al. (2014) A reference requirements set for public service provision enterprise architectures. Softw Syst Model 13(3): 991–1013. doi: 10.1007/s10270-012-0303-7

28. Califf C, Lin X, Sarker S (2012) Understanding Energy Informatics: A Gestalt-Fit Perspective. In: Joshi KD, Yoo Y (eds) AMCIS 2012 Proceedings, p 13
29. González J, Appelrath H (2010) Energie-RMK" - ein Referenzmodellkatalog für die Energiewirtschaft. In:
30. Lampropoulos I, Vanalme GMA, Kling WL (2010) A methodology for modeling the behavior of electricity prosumers within the smart grid. In: IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe), pp 1–8
31. Eßer A, Franke M, Kamper A et al. (2007) Future power markets. Wirtsch. Inform. 49(5): 335–341. doi: 10.1007/s11576-007-0077-7
32. Reichert A, Otto B, Österle H (2013) A Reference Process Model for Master Data Management. In: Alt R, Franczyk B (eds) Proceedings of the 11th International Conference on Wirtschaftsinformatik (WI2013), pp 817–845
33. Otto B, Ofner MH (2010) Towards a Process Reference Model for Information Supply Chain Management. In: ECIS Proceedings, p 75
34. Wieringa R (2009) Design science as nested problem solving. In: Vaishanvi V, Purao S (eds) the 4th International Conference on Design Science Research in Information Systems and Technology (DESRIST '09), p 8
35. Wieringa R, Moralı A (2012) Technical Action Research as a Validation Method in Information Systems Design Science. In: Peffers K, Rothenberger M, Kuechler B (eds) Design Science Research in Information Systems. Advances in Theory and Practice, vol 7286. Springer Berlin Heidelberg, pp 220–238
36. Becker J, Delfmann P, Knackstedt R et al. (2002) Konfigurative Referenzmodellierung. In: Becker J (ed) Wissensmanagement mit Referenzmodellen: Konzepte für die Anwendungssystem- und Organisationsgestaltung ; mit 13 Tabellen. Physica-Verl., Heidelberg, pp 25–144
37. Runeson P, Höst M (2009) Guidelines for conducting and reporting case study research in software engineering. Empir Software Eng 14(2): 131–164. doi: 10.1007/s10664-008-9102-8
38. Timm F, Köpp C, Wißotzki M (2015) Initial Experiences in Developing a Reference Enterprise Architecture for Small and Medium-Sized Utilities. In: Espana S, Ralyte J, Soffer P et al. (eds) PoEM 2015 Short and Doctoral Consortium Papers, Valencia, pp 31–40
39. Fettke P (2014) Eine Methode zur induktiven Entwicklung von Referenzmodellen. In: Kundisch D, Suhl L, Beckmann L (eds) MKWI 2014 - Multikonferenz Wirtschaftsinformatik : 26. - 28. Februar 2014 in Paderborn, Paderborn
40. Stirna J, Persson A, Sandkuhl K (2007) Participative Enterprise Modeling: Experiences and Recommendations. In: Krogstie J (ed) Advanced information systems engineering: 19th international conference, CAiSE 2007, Trondheim, Norway, June 11 - 15, 2007 ; proceedings, vol 4495. Springer, Berlin u.a., pp 546–560
41. Gregor S, Hevner AR (2013) Positioning and Presenting Design Science Research for Maximum Impact. MIS Q 37(2): 337–356
42. Goldkuhl G, Lind M, Seigerroth U (1998) Method integration: The need for a learning perspective. IEE Proc., Softw. 145(4): 113. doi: 10.1049/ip-sen:19982197

Enterprise Modeling Support for SOA Migration

Sybren de Kinderen and Monika Kaczmarek-Heß

Information Systems and Enterprise Modelling, University of Duisburg-Essen, Essen, Germany
{sybren.dekinderen,monika.kaczmarek}@uni-due.de

Abstract. The migration to a Service Oriented Architecture (SOA) is a complex undertaking in terms of aligning business and IT concerns as well as analysis of technical aspects. Conceptual modeling can be helpful for supporting SOA migration by (1) bridging the gap between business and IT concerns, and (2) analyzing the as-is and to-be IT infrastructures. We contribute language requirements derived from SOA migration literature, and extend an IT infrastructure Modeling Language, ITML, to support SOA migration. We illustrate the extended ITML with a documented SOA migration case.

Keywords: enterprise modeling, service-orientation, SOA migration

1 Introduction

Since its emergence and subsequent mainstream acceptance during the early 2000s [1, p. 75], adopting a Service Oriented Architecture (SOA) remains an important concern for many organizations [2, 3]. Service orientation (ideally) fosters organizational flexibility and agility by promoting (1) re-use of the functionality offered by services [5], (2) modifiability of (IT) functionality, which is realized by loose coupling [6, p. 64].

Successfully migrating an organization towards service orientation has however proven challenging [7, 8]. Conceptual modeling has the potential to play an important role in supporting SOA migration, in particular to foster (1) communication between business experts and IT experts [10], so as to ensure that IT functionality is driven by business concerns and vice versa; (2) to perform an in-depth analysis of the as-is and to-be IT elements [9, 11] so as to, e.g., identify functionality of legacy systems in need of wrapping [11]. Although various modeling languages exist that allow for expressing service orientation from various angles (cf. [12–14]), these languages often on purpose forgo the level of detail that is required in analysis of an IT infrastructure for the needs of SOA.

In order to address this gap, we focus on the following question: *What should be the scope and characteristics of a modeling language able to support SOA migration projects?* To address it, we first identify a set of requirements based on the analysis of SOA migration literature and of characteristics of SOA migration projects and then, extend an already existing modeling language, namely ITML [4], to account for missing aspects. The proposed extensions have been evaluated against the requirements and with an extensively documented SOA example.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

de Kinderen, S.; Kaczmarek-Heß, M. (2017): Enterprise Modeling Support for SOA Migration,
in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung
Wirtschaftsinformatik (WI 2017), St. Gallen, S. 346-349

2 Modeling Support for SOA Migration: Requirements

'Migration' is usually understood as moving an existing operational system to a new technological or computing platform, while retaining the data and functionalities of the moved system (cf. [15]). SOA migration projects requires knowing at least the as-is and to-be states of the IT infrastructure and how these support business concerns (cf. [9, 11]).

Table 1. Requirements on a conceptual modeling language supporting SOA migration

RQ	*Requirement and Candidate Concepts*
	I Requirements For As-Is Models
1	The modeling language should allow for expressing IT landscape elements. Candidate concepts: Database, Database management system, Middleware, Server
2	The modeling language should allow for expressing the dependencies between IT landscape elements. Candidate concepts: uses, provides, runs on
3	The modeling language should account for non-functional properties of legacy systems. Candidate concepts: mission criticality, source code availability, impl. lang., code complexity
	II Requirements For To-Be Models
4	A modeling language should provide dedicated concepts that allow to model a service and its relevant types. Candidate concepts: Service, WebService, Interface
5	The modeling language should allow for relating a service to its underlying implementation, in accordance with the migration strategy. Candidate concepts: Wrapper, provides, runs on, uses
6	The modeling language should account for quality attributes of service oriented concepts. Candidate concepts: various QoS characteristics
	Overall Requirements
7	The modeling language should allow for expressing dependencies between the IT landscape and the organization action system. Candidate concepts: supports, context of use

The as-is state usually encompasses information about legacy systems. Legacy systems are systems that are usually hard to modify and expensive to maintain. However, at the same time these systems are often mission-critical and thus, must be operational at all times [9]. The to-be state reflects the service oriented design of the architecture. As understanding the as-is and to-be states is important for carrying out the migration, the aim of the modeling language should be to: (1) provide knowledge on the current state of the IT infrastructure with the focus on legacy systems (cf. RQ1 & RQ3) and (2) express the to-be state of the service orientation and reflect the changes that should be performed following the selected migration strategy (cf. RQ4).

For the as-is and to-be state IT infrastructure, we are interested in expressing the observable functionality of IT infrastructure elements, which translates into the requirements: *IT infrastructure elements* (RQ1) and their *interdependencies* (RQ2).

Furthermore, for any meaningful analysis of the possible behavior of IT infrastructure elements we need to analyze non-functional attributes [11] (RQ3).

Concerning the to-be IT landscape, the language should provide (rudimentary) expressiveness for service orientation. This entails to (1) express services and related subtypes, in addition to their quality attributes, through the language (RQ4), as well as to (2) relate these services to elements of the IT infrastructure (RQ5). The latter relation is important for the aim of the language: to analyze how an IT infrastructure should be changed to realize the functionality offered by a service.

In addition, SOA migration should be considered from both an IT infrastructure and a business perspective (cf. RQ7). Business processes largely drive what is implemented in terms of IT support [8, 16], and vice versa.

3 Extended ITML

We now briefly illustrate SOA extensions of our language, called the extended IT Modeling Language (ITML), which is based on [4]. For illustration purposes, in Fig. 1 we modeled a to-be SOA of the ACME insurance company [6, pp. 541–578]. We focus on three features of extended ITML, illustrated by the labels 1-3 (cf. Fig. 1).

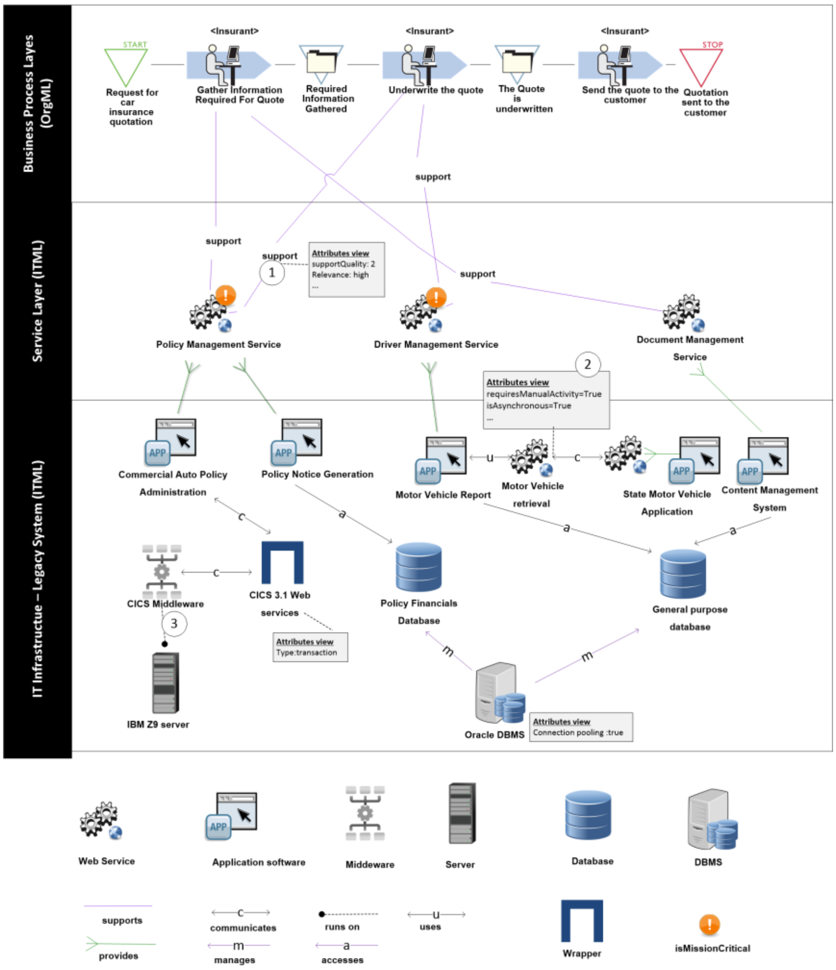


Figure 1 ACME insurance's desired service orientation, modeled in ITML

Extended ITML allows for (1) expressing, when used in tandem with the other MEMO languages [17], how IT functionality supports business processes (cf. RQ7). E.g., the web services “policy management service” and “driver management service” both support – Label 1 – the business process “underwrite the quote” (with *relevance:high*), (2) expressing non-functional attributes (RQ3&RQ6), such as – Label 2 – the necessity of a web service to support asynchronous communication, as well as a wrapper supporting bi-directional communication, and (3) inventorying relevant IT infrastructure assets and their relations (cf. RQ1&RQ2&RQ5), such as – Label 3 - the desire to keep using legacy bulk transaction processing functionality via a wrapper.

References

1. Erl, T.: Service-oriented architecture: concepts, technology, and design. Pears. Edu. (2005)
2. MacLennan, E., Van Belle, J.P.: Factors affecting the organizational adoption of service-oriented architecture (SOA). ISeB **12**(1) 71–100 (2014)
3. Alwadain, A., Fielt, E., Korthaus, A., Rosemann, M.: Empirical insights into the development of a service-oriented enterprise architecture. DKE **105** 39-52 (2016)
4. Heise, D.: Unternehmensmodell-basiertes IT-Kostenmanagement als Bestandteil eines integrativen IT-Controllings. Logos, Berlin (2013)
5. Razavian, M., Lago, P.: Towards a conceptual framework for legacy to SOA migration. In: ICSOC/ServiceWave Workshops, pp. 445–455. Springer (2010)
6. Rosen, M., Lublinsky, B., Smith, K.T., Balcer, M.J.: Applied SOA: Service- Oriented Architecture and Design Strategies. Wiley Publishing (2008)
7. Hirschheim, R., Welke, R.J., Schwarz, A.: Service-oriented architecture: Myths, realities, and a maturity model. MIS Quarterly Executive **9**(1) (2010)
8. Rabelo, R.J., Noran, O., Bernus, P.: Towards the next generation service oriented enterprise architecture. In: IEEE 19th Int’l EDOC Workshop, pp. 91–100. IEEE (2015)
9. Khadka, R., Saeidi, A., Jansen, S., Hage, J..: Migrating a large scale legacy application to SOA: Challenges and lessons learned. In: WCRE 2013. pp. 425–432. IEEE (2013)
10. Razavian, M., Gordijn, J.: Consonance between economic and it services: finding the balance between conflicting requirements. In: REFSQ, pp. 148–163. Springer (2015)
11. Razavian, M., Lago, P.: A systematic literature review on SOA migration. Journal of Software: Evolution and Process **27**(5) 337–372 (2015)
12. The Object Management Group: Service oriented architecture Modeling Language (SoaML), version 1.0.1 (2012)
13. Terlouw, L.I., Albani, A.: An enterprise ontology-based approach to service specification. IEEE Transactions on Services Computing **6**(1) 89–101 (2013)
14. The Open Group: ArchiMate 2.1 Specification: Open Group Standard. The Open Group Series. Van Haren, Zaltbommel (2013)
15. Bisbal, J., Lawless, D., Wu, B., Grimson, J.: Legacy information systems: Issues and directions. IEEE Softw. 16(5) 103–111 (1999)
16. Papazoglou, M.P., Van Den Heuvel, W.J.: Service-oriented design and development methodology. Int J Web Eng Tech **2**(4) 412–442 (2006)
17. Frank, U.: The MEMO Meta modeling Language (MML) and Language Architecture. 2nd Edition. ICB-Research Report 43, University of Duisburg-Essen (2011)

Development and Test of a Semi-structured Explorative Survey Methodology to Analyze Appropriate Learning Methods for Technology-related Training across the Phases of Technology Use

Diana Renner[1] and Sven Laumer[1]

[1] University of Bamberg, Department of Information Systems and Services, Bamberg, Germany
{diana.renner,sven.laumer}@uni-bamberg.de

Abstract. Training is regarded as an indicator for organizations' performance. To analyze appropriate learning methods for technology-related training across the phases of technology use we develop a semi-structured, explorative survey methodology. The methodology is tested with a sample of 53 healthcare professionals from China, France and the USA. Based on that, lessons learnt are concluded that confirm the usefulness of the methodology for research and practice. In addition, further improvements and enlargements of the methodology are illustrated in this paper as basis for further research.

Keywords: blended learning, training needs analysis, healthcare, technology, technology use

1 Introduction

"*Extremely powerful*" and "*cost-effective-investment*" [1, p. 147] – these two quotes relate to the characteristics of training. In the context of corporate education, training is regarded as an indicator for increasing the organization's performance [1-2]. Thus, studies focused on the investigation of pre-training, training and post-training activities [3]. From a general training research perspective, we know that the training processes start with the analysis of training needs followed by the design and delivery of trainings and ends with the evaluation and transfer of it [4]. The starting process is often referred to a training needs analysis (TNA) and concentrates on the assessment and analysis of data to determine training needs for an organization [5]. Necessary skills and knowledge that need to be acquired by the workforce are analyzed to ensure the organization's performance [2].

The appropriateness of learning methods for the delivery of knowledge and skills is not analyzed in the current practice of TNA. This confirms the request for research of Santhanam et al. [3] regarding how to match appropriate learning methods with characteristics of the trainee to ensure positive learning outcomes. The achievement of this match represents a challenge especially for blended learning concepts. These

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Renner, D.; Laumer, S. (2017): Development and Test of a Semi-structured Explorative Survey Methodology to Analyze Appropriate Learning Methods for Technology-related Training across the Phases of Technology Use, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 350-364

concepts are characterized through the combination of traditional learning methods such as classroom training and online learning methods such as web-based trainings [6]. During the last years, blended learning is increasingly used in organizations for training delivery [7] and requires a decision whether rather online or offline learning methods for the delivery of knowledge and skills should be used.

Furthermore, when a new technology is implemented the knowledge and skills which need to be trained are defined by the project management or general management [8]. The analyzed training needs are often short-term and task-oriented; however, the use of technology covers multiple years and is long-term oriented. Therefore, several phases of technology use are differentiated [9], which indicate that across all phases of technology use training needs arise and trainings are delivered; however, the respective needs for each phase are different [10].

Next to the skills and knowledge needed to be trained, the environment of the learner as well as his/her motivation is important to result in a good performance [11]. Therefore, we can assume that if the individual feels the training method as appropriate, his/her motivation increases and leads in conclusion to a better performance. Technology-related training should address the training needs of the individual for the relevant phase of technology use with regards to relevant knowledge and skills but also consider the appropriateness of learning methods based on the characteristics of the individual trainee such as working environment, training time, etc.

This aspect is not covered in the current state of TNA such that no methodological approach exists that focuses on the analysis of appropriate learning methods for technology-related training across different phases of technology use. Especially the appropriateness of blended learning for technology-related training and thus the appropriate interplay of traditional learning and online learning methods over a longer timeframe and in particular across the different phases of technology use remain unknown. Therefore, we address this gap and focus on the following research question:

> *How to analyze appropriate blended learning methods for technology-related training across the phases of technology use?*

To answer this research question, we develop a semi-structured, explorative survey methodology based on TNA research. We applied this methodology within a study in the healthcare sector and interviewed 53 subjects from ten institutions located in China, USA and France. Within our paper we will introduce the semi-structured, explorative survey methodology in detail and discuss the lessons learnt from its initial use with 53 employees. Therefore, we follow the schema of Gregor and Hevner [12] but focus on the design of the artifact and its first application in practice within this paper. The conducted evaluation investigates the usefulness and comprehensibility of the methodology for practical application, as it is important in a first step to develop a methodology that is accepted by learners [11]. It is shown that the positive evaluation of a learning method by the learner him/herself leads to a positive increase of his/her learning outcome [13]. In a second step, the designed artifact needs further be evaluated according to design science guidelines [12], [14], [15] to demonstrate that by using the suggested methodology to select an appropriate mix of learning methods for technology-related training across the phases of technology use will improve the learning outcomes. However, this part of the evaluation is not part of this paper.

We structure the paper as followed. We focus in the theoretical background on TNA and provide an overview about the state-of-the art. In addition, we explain the concept blended learning and the different phases of technology use. Then the development and evaluation of the methodology follow. Based on our first application, we will present our lessons learnt and how we adjusted the methodology. Implications for theory and practice will be provided in the discussion followed by future research and the limitations of the paper.

2 Research background

We introduce the TNA as start point of training processes in the following. In addition, an overview about current research of TNA is illustrated. Afterwards, blended learning and the phases of technology use are explained.

2.1 Training needs analysis

In general, a training need analysis is defined as *"process of gathering, assessing and analyzing data to determine the training needs for an organization"* [5, pp. 393-394]. The goal of the TNA is the identification of training needs as well as the design and development of ways and resources to address and satisfy the needs by the most possible cost-effective and efficient manner. Based on an organization's current and desired performance levels, data will be collected and interpreted by using methodical investigations and analyses [1-2]. The TNA consists of the organizational analysis, the task and knowledge, skill, and ability analysis and the person analysis. The organizational analysis focuses on the short- and long-term goals of the organization as well as on the trends that might affect these goals. The task and knowledge, skill, and ability (KSA) analysis concentrates on the analysis of the job that needs to be performed by the trainees upon completion of the training program. The third part of the TNA, the person analysis, investigates how a specific employee is carrying out the tasks of his/her job [16]. The data collection of the TNA can be done through different methods such as surveys/questionnaires, interviews, observations etc. [8]. The collected data will be interpreted and provide the basis for further activities such as training delivery.

The results of our literature analysis[1] indicate that only the task and KSA analysis are conducted within TNA reported in the literature [17-19]. Moreover, apart from Waldman et el. [20] who use interviews and survey in their studies, the other identified research studies mainly use surveys as method for performing the TNA.

[1] To get an overview about past research we screened the databases of the IS senior basket of eight and relevant journals of human resources and the organizational context (journal rated A+ and A) as well as general economics (A+)). We used the vhb-journal ranking (http://vhbonline.org/VHB4you/jourqual/vhb-jourqual-3/) as orientation. We used "training needs", "training needs analysis" and "training needs assessment" as search terms.

The literature review of Moore and Dutton [21] states that the implementation of TNA is not done in the way as it is suggested by training theorists. Although the literature review is long time ago, less research focuses on TNA in the following years such that their conclusion is still valid. Ford and Noe [18] identify in their study high impact of the position of managers in the organization and small impact of individual characteristics such as the attitude towards the utility of training on the self-assessment of training needs by managers. Phang et al. [19] develop a TNA module to support employees in their selection of the right training modules based on their analysis of training needs. The current state of the TNA research indicates that TNA covers a small part in the research field of training processes [4]. The analysis of training needs concentrates on skills and tasks of jobs and in conclusion, on relevant training content. In addition, the analysis has no temporal focus such as milestones or phases of use. Hence, we can conclude from prior research that there is a gap in the literature that focuses on how to analyze the appropriateness of learning methods and especially the mix of online and traditional methods across the different phases of technology use. Therefore, we will focus on this gap in this paper by proposing a TNA method with a focus on blended learning methods such that we introduce in the following blended learning and the phases of technology use.

2.2 Blended learning

In general, learning methods can be classified in traditional learning methods and online learning or electronic training methods [11]. Both methods are characterized through their benefits and drawbacks. Traditional learning methods have a high degree of interactivity which in contrary is often related to high costs. The benefits for the employees of being out of their daily working routine e.g. for a classroom training and having the instructor directly available for comments and questions are connected to missing working time and possible costs for traveling [7]. Online learning methods that are also known as e-learning offer the employees flexibility regarding the aspect when and where to learn and how much to learn. The instruction is often asynchronous as online learning methods are delivered by a CD-ROM, the internet or the intranet. Instruction can also take place synchronous e.g. in form of a virtual classroom training or remote training [22]. The use of a learning management system (LMS) enables to provide learning material without time and place-restrictions. LMS integrate features such as online discussion forums for the exchange of the learners or further tools such as simulation which can be used for training issues [23]. Moreover, it enables the organization to track the learning outcomes and progress of its employees [7], [22]. This aspect is often related to negative aspects of online learning such as the high front-up costs for the organization. In addition, employees often complain about the lack of interaction that might result in demotivation and feel isolated from the other learners [7]. The implementation of blended learning – which reflects the approach of combining traditional and online learning methods – tries to avoid the negative aspects and benefit from the positive aspects of both learning methods [6-7].

The results of studies indicate that the blended learning approach leads to better learning outcomes compared to only traditional learning [24] or online learning [25]. Learning outcomes can be differentiated into affective, cognitive and skill-based outcomes [26]. Previous research shows that blended learning has a positive influence on affective outcomes such as the satisfaction of the learner or cognitive outcomes such as the effective knowledge transfer [24]. In addition, past studies in the blended learning context mainly focus on comparisons between blended learning and traditional learning concepts or online learning concepts. Potential predictors for course outcomes are examined [27]. In the healthcare IS research contextual factors which have impact on the relationship between blended learning and positive learning outcomes are investigated. The studies of Buyl and Nyssen [28] and Lopez-Campos et al. [29] examine that based on the user group different blended learning concepts are necessary to ensure positive learning outcomes. In addition, culture as contextual factor is examined by the Sánchez-Mendiola et al. [30]. They confirmed that the concept of blended learning is especially successful in developing countries. In conclusion, there is no research which uses blended learning concretely in the context of technology-related training. In addition, no research focuses on the contextual factor time such that designing blended learning concepts over a longer timeframe. We use the examined contextual factors as basis for the development of appropriate blended learning concepts as they are important in the context of technology use [10] as it will be discussed in the following and hence, also for designing a blended learning concept for technology-related training.

2.3 The different phases of technology use

In general, the phases of technology use cover six phases of a technology in connection to its diffusion within the user groups [31]. The first phase, *pre-implementation*, can be defined as the period before the new technology is available for use in an organization [31] . This phase includes the process of defining, creating, and obtaining the tools, documentation, procedures, facilities, and any other physical and informational resources which are done before the implementation of the technology takes place [32]. From a training perspective the focus of the first phase is especially on making the users ready for the implementation phase. In the *implementation* phase, which is the period when the new technology has been implemented and employees are starting to use it to the point when the technology is routinely used in the daily working [33], employees get in touch with the new technology and start using it. Therefore, users need to be supported in using the technology during the implementation phase. Follow-up challenges need to be addressed in the following *follow-up* phase. Finally, in the post-implementation period, which is the period when the technology is becoming or has become part of the organization and employees use it in a routine way [34], the use of the technology itself or of specific features need to be refreshed, new employees need to be trained in using the technology or new features need to be introduced and learned when small changes or updates of the technology are implemented [32]. As a result, we differentiate between three further phases *refresh*, *new hire* and *new feature*. In summary, there are six phases of technology use which can be distinguished as illustrated by Figure 1 (left side). All the phases result in different training requirements. We will use these six different phases as a basis to reveal the appropriateness of traditional and online training methods in each phase.

3 Development of methodology

Summarizing the findings of our research background we conclude that there is no methodology covering the analysis of the appropriateness of learning methods in general and for blended learning in particular. In addition, we found no methodology that analyzes appropriate methods for technology-related training across the phases of technology use. To develop the required methodology, we use a design science research approach and followed the guidelines of Hevner et al. [14] and Gregor and Hevner [12] to design our methodology artifact.

The need for the design of an artifact results from a business problem located in the healthcare sector. The company produces healthcare technology and related services such as training in order to use the technology effectively and efficiently in the daily working routine. The job group of healthcare professionals is characterized through a high specialization in their job and a wide array of tasks performed by the group each working day [35]. The tasks are related to different skills and knowledge, expertise and former education implicates differences with regards to the use of for example healthcare information systems by nurses and physicians. The high responsibility related to the job of healthcare professionals results in a high training priority of the healthcare workforce to ensure the appropriateness of treatments and, in conclusion, the patients' health. Therefore, the use of technology is crucial in the healthcare context. As the job responsibilities and tasks vary between nurses, physicians, technicians etc., healthcare information systems are customized to cover the specialized need of the different user groups [36]. Conversely, this means that the customization of training to meet the training needs of different groups can be assumed. As a result technology-related training is of high importance for the healthcare technology company. To solve the business problem and to be able to work out concrete planning of training delivery over a longer period as well as possible investments for tools and infrastructure, the management requires a methodology for analyzing technology-related training needs which focuses on how to deliver technology-related training. The content delivered by the trainings (with regards to tasks and related knowledge and skills which need to be learned by the healthcare professionals) is well-defined. As the knowledge base of IS research [14] includes currently no methodologies facing this business need (see section 2), the development of an artifact is necessary. To get the full picture of the IS research framework we supplement the analysis of the knowledge base by the investigation of the environment in which the designed artifact should be applied. With the help of the analysis we want to get an overview about the current state-of-the art and to draw conclusions for the design of the artifact. Therefore, we analyze in the first step all relevant information from customers of the healthcare technology company with regards to methods used to analyze training needs. This provides us a first overview about the current methods used for technology-related training, issues and improvements. In the second step, we examine internally which training methods are currently used or in planning for technology-related training by conducting a survey with product managers who are responsible for the creation and implementation of training strategies. In the third step, we analyze the training offerings of further healthcare technology companies. Based on the three-step-analysis, we conclude that often the management of healthcare professionals is asked about training needs and related issues and not the end user of the technology. The feedback indicates that training needs exists also after the implementation of the technology for refreshing knowledge, extending expertise or the training of new employees. For training delivery, product managers mainly decide to use traditional learning methods such as on-site training, which takes place directly at the institution of healthcare professionals when the implementation of the technology takes place. On-site training is sometimes combined with web-based training or delivered by CD/DVD dependent of the infrastructure of the institutions. The training delivery focuses on the phases *implementation* and *follow-up*. Besides, healthcare professionals feel "over surveyed", as they commented to be too much contacted for surveys for all different cases and they indicated that they only want to participate in surveys they evaluate as positively for themselves.

Hence, based on this analysis, we designed a methodology that enables to analyze appropriate learning methods for technology-related training across the phases of technology use. A first evaluation of the artifact will be done. The evaluation focuses on the acceptance of the new methodology by learners as our analysis indicates that healthcare professionals feel "over surveyed" and that they only want to participate in surveys they evaluate as positively for themselves.

4 Semi-structured, explorative survey methodology

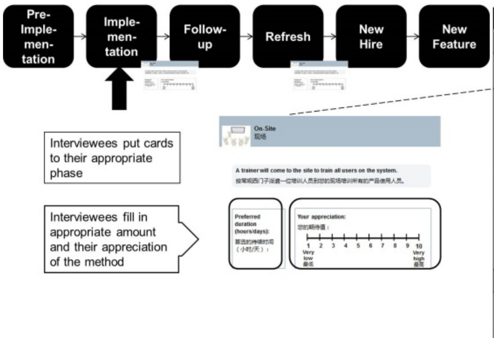
To assess the appropriate mix of traditional and online learning methods that addresses the training needs of individuals for the relevant phase of technology use we design a new methodology artifact. The proposed semi-structured, explorative survey methodology consists of two parts.

The first part focuses on general information with regards to the job, training and product experience of the interviewees. The interviewees are asked how they would classify their user role (*regular user* who uses the basics of the technology, *expert users* who are key users and have deep knowledge, and *head* users who are in a management role and often formerly worked as expert user). This enables the interviewer to get a first profile of the healthcare professional.

The second part focuses on the appropriateness of different training methods across the different phases of technology use. Therefore, we provided a list of different training methods and tools that can be used in the learning process (see Figure 1). We differentiated between traditional and online training methods and described each method in detail. This classification is based on Blanchard et al. [11]. In addition, we introduce the six phases of technology use as described in section 2. The data collection focuses on evaluating the appropriateness of different learning methods and on the amount and appreciation of the learning method. We regard the appropriateness of a training method or tool as suitable and compatible for individual subject to solve his/her issue of training needs for the phase of technology use [17]. The amount of the training method is simply the amount of time that the individual likes to spend with the training method or tool such as hours or days and is important for the planning of training [3], [17]. The appreciation considers *"a meaningful response associated with mixed emotion"* [37, p. 398] of the subject with regards to the training method or tool. Therefore, the relative value of the method or tool for the individual is in focus which covers also an important affective measure in training research [3], [17].

To **collect the data**, we suggest using a Din A1 poster and cards as additional guideline. The poster shows an overview about the six phases of technology use (see Figure 1). With the help of the cards the different learning methods are shortly explained (see the overview in Figure 1) and the amount of the training method or tool can be filled in as well as the appreciation of the individual regarding the training method or tool can be specified. For additional suggestions of training methods by interviewees empty cards are provided which can be filled out by the interviewees as well. The procedure of the data collection is suggested as follows: (1) The interviewer explains the different phases of technology use to the interviewee; (2) the interviewer explains the cards to the interviewee; (3) the interviewee choses his/her appropriate learning methods and tools for each of the phases of technology use and put the cards to the relevant phase, (4) the interviewee filled out the cards and specified the amount of training as well as the appreciation of the training method on a scale from 1 (very low) up to 10 (very high) in detail, (5) each interviewee is asked for a justification, why s/he has laid down the learning method to one of the phases of technology use, why the specified amount of the learning method is needed and why the learning method is more or less appreciated. We suggest to record the survey and that the interviewer should take notes. Figure 1 illustrates the procedure of this interactive part.

For **analyzing the data** we suggest that on the one hand one should evaluate the notes taken by the interviewer. For better comparison, the notes are structured according to comments to the phases of technology use, training methods, amount or appreciation and further comments. In addition, for clarifying or analyzing some aspects in detail, the recorded tapes can be used. On the other hand, the cards chosen by the subjects are evaluated. Thus, one regards the appropriateness of learning methods in relation to the phases of technology use in general. Therefore, one examines the overall number of cards laid down by each interviewee to all phases of technology use. In addition, one counts the cards put down to each phase of technology use. This provides a first overview about the interviewees' appropriateness of some training methods for different phases. Moreover, it shows the extent of training needs for each phase. If the interviewees wrote down additional notes on cards one can listen again to the tape to have a full understanding as the interviewee often provide more information regarding his/her notes by explaining it to the interviewer and/or translator. Furthermore, one can analyze descriptive data by the percentage of traditional learning and online learning methods in general and in detail for each of the six phases. In addition, one can investigate the percentage of tools chosen for supporting the traditional and online learning methods. The amount of training methods can be evaluated by one in the first step per hour; in the second step one can calculate the mean of each training method for each of the six phases of technology use. The appreciation of the training methods can be evaluated in general by analyzing the mean of the training method independent of the phases of technology use and in addition separated for each phase of technology use.



Method	Explanation
Traditional learning	
On-site	A trainer will come to the site to train all users on the system.
Classroom training	System users from different facilities are trained at a training center (e.g. of a vendor) to get more routine or more expertise on the system.
Clinical workshop	System users from different facilities meet at a renowned site/hospital to be trained by an advanced system user for more expertise.
Symposium	During a scientific conference, attendees can participate at various sessions around all modalities.
Consulting	A specialist will come to consult the staff of the site on different topics e.g. image quality or utilization.
Train the trainer	Course at a training center (e.g. of a vendor) that focuses on didactical skills for the internal training of customers.
Online learning	
Virtual classroom	Online meeting space for real-time learning activities with a trainer and participants from other sites/hospitals.
Web-based training	Interactive learning program on the user's computer with internet and browser access, no trainer.
CD/DVD	Interactive learning program on the user's computer, no trainer.
Podcasts	Audio or movie that can be played on mobile devices or computer (e.g. a clinical speaker informs users about best practice of improving image quality).
Remote training	Interactive one-on-one course with a remote trainer. Via remote access training can be directly conducted on the system.
Remote assistance	To receive immediate hotline/online application support in case a user gets lost during an examination.

Tool	Explanation
Online discussion forum	A forum that is accessible for system users for questions and exchange (best practice). Experts take care of the forum.
Online help	Online help provides topic-oriented, referenced information on the software. It is intended as user assistance.
Mobile simulator	An online information portal (e.g. of a vendor) that offers e.g. scan simulation. Different protocols and functionalities can be tested.

Figure 1. Proceeding of interactive part and overview about learning methods and tools for supporting

Hence, the proposed methodology is characterized as interactive. The interviewee is involved directly by a task of selecting appropriate traditional and online learning methods. The interviewer can provide guidance to the interviewee if s/he has questions. In addition, the interviewer can directly ask the interviewee why s/he has selected this learning method and why this amount is needed as well as why the method is more or less appreciated by the interviewee. This results in the detailed explanation of the appropriateness of training methods across the phases of technology use. To test the appropriateness of the methodology we evaluated it in a study in the healthcare context as it will be described next.

5 Evaluation of semi-structured, explorative survey methodology

To evaluate the usefulness, the practicality and comprehensibility of the designed methodology we conduct a study. The objective of the evaluation is the first presentation of the methodology and its practical application. A full evaluation of the design artifact has to be done according to the guidelines of design science research in future research to concretely demonstrate the utility, validity, quality and efficacy of the artifact [12], [14], [15]. The study includes three large Chinese hospitals, three large French hospitals and one large diagnostic imaging center in France as well as three large hospitals in the USA. The contact to the institutions was established by the healthcare technology company that highlighted the need for developing a new methodology for assessing the appropriateness of training methods.

In total, 53 Healthcare professionals were interviewed using the newly developed methodology and welcomed to provide feedback to the methodology. The average age of interviewees is 38.5 years (seven subjects provide no data about their age). 24 were female and 29 were male interviewees. An overview about our sample is illustrated in Table 1. The interviews took between 45-60 minutes. We select our interview partners in dependency of the technology they use. We ensure that that all interview partners work in radiography departments such that they use technology systems like angiography or x-ray (AX), low complexity of technology, computed tomography (CT), medium complexity, and magnetic resonance imaging (MRI), high complexity. Moreover, we select our interview partners according to three user roles: regular user, expert user and head.

Table 1. Overview of interviewees

		User role		
		Regular	Expert	Head
Technology	AX	CN (2); FR (0); USA (0)	CN (2); FR (3); USA (0)	CN (2); FR (1); USA (0)
	CT	CN (1); FR (1); USA (5)	CN (2); FR (6); USA (5)	CN (1); FR (1); USA (5)
	MRI	CN (1); FR (0); USA (1)	CN (2); FR (0); USA (0)	CN (2); FR (1); USA (1)
CN: China; FR: France; brackets cover number of interviewees				

For collecting our data we interviewed each professional first using the proposed method and afterwards we asked them about their evaluation of the interview conducted. For the second part, we asked questions such as *"What is your opinion about the methodology?" "How would you rate the usefulness of the method?" "Would you use the methodology in the future?"*

In addition, we observed the subjects while conducting the method. We analyzed their reactions regarding the proceeding, their tasks and if the method and its proceeding and outcome were understandable. We made additional notes if the interviewees asked questions regarding the method. We classified the notes taken by the interviewees according to questions about the proceeding in general, the comprehensibility of the method, improvements for the method and further notes.

We coded the feedback based on the transcripts of the second part of the interview and we coded the emotions of the interviewees by conducting the methodology as well as the reaction of the interviewees for providing feedback. We classified the emotions in positive (e.g. excited, satisfied), negative (e.g. bored, distracted) and neutral. Hence, we identified incidents that the interviewed healthcare professionals rather evaluate positively or negatively. Furthermore, we presented our methodology and the results of the first application to the management of the healthcare technology company. Based on the feedback and the analysis of the interviews we deduced lessons learnt from the first application for further adjustments of the methodology which will be explained in the next section.

6 Lessons learnt from first application

As a result of the demonstration and first evaluation of our designed artifact [14] we conclude four lessons learnt addressing the practical applicability of the methodology:

Lesson learnt 1: Interactivity takes time but it is worth. We estimated for conducting the survey - including all parts - 30 to 40 minutes. Whereas the first part about questions regarding the experience of interviewees was done efficiently, the second part took time because of the explanation of the additional material such as the poster and the different cards. In China and France where the interviewers were accompanied by a translator, it takes more time for translations even when the interviewees clarified for example their understanding of the training methods or tools to ensure the correctness of their selected cards. In some cases, the interviewees did not finish the survey despite lack of time.

In general, the feedback of the interviewees to the semi-structured, explorative survey methodology was positive. The interviewees felt satisfied as the methodology enables them to plan their ideal blended learning concept for learning a new technology. One American interviewee said: "*We had some time ago the implementation of a new technology. So I could reflect about the training we had and if something was missing and could be improved the next time.*" (male, 41, expert user CT). In contrary to the French and U.S. interviewees, the reactions of Chinese interviewees were more characterized through shyness and restraint. Therefore, more guidance from the interviewer and the translator was necessary. The feedback to the survey was also positive. As a result, the methodology is in general well applicable for practical use but should be adjusted regarding the aspect planning. Enough time for the performance should be calculated and also communicated to the interviewees in advance to ensure that the interviewees have enough time to complete the survey.

Lessons learnt 2: Adaption of the different phases of technology use and the methods based on the context of technology. We applied the survey methodology in the healthcare context. Therefore, other training methods can be more appropriate in other business contexts. In addition, the phases of technology use can be different based on the characteristics and use of the technology. In our study, 52 of the interviewees agreed with the six phases of technology use. Only one Chinese interviewee provided feedback that she would add another phase named "improve clinical experience" after the sixth phase. This phase is *"for improving the clinical expertise of the user and to extent the clinical knowledge"* she explained (female, 39, MRI expert user). With regards to the training methods, additional methods to the cards provided were suggested by the interviewees. Four interviewees from France proposed for example the "visit of another institution" as another training method to the *pre-implementation* phase. We can summarize that independent of the technology the interviewees are using; further methods and tools may be appropriated additional to the setting of our methodology. Therefore, the possibility for additional proposals by the interviewees should be included in the survey methodology.

Lessons learnt 3: Estimation of training amount as challenge. In the methodology we ask the interviewees also for the appropriate amount of the selected method/tool. Based on our first experience in using the survey methodology, we can conclude that this aspect is difficult to specify for the interviewees especially the amount for spending time with the tools. Therefore, we focus on guiding the interviewees to conclude what works best and how our methodology might be improved. As a result, the interviewees took additional notes to the appropriate amount they like to spend with the tool or method. The American MRI head user selected the tool mobile simulator for the pre-implementation phase of technology use but added *"it would be good to have the mobile simulator around three months in the pre- and post-install phase available to support internal training"* (male, 41). The additional information provided by the interviewees facilitates on the one hand the planning of training and the possibility to customize training according to special needs. On the other hand, comparisons between countries, for example, are difficult.

Lessons learnt 4: Expanding the horizon – for business and research. The semi-structured, explorative survey methodology and a summary about important results were also presented to the management of the healthcare technology company. With regards to the survey methodology the feedback of the management was positive to solve the business problem and for practical application. One manager commented *"A really interesting and especially innovative approach. […] This enables us to work on the customization of training needs […] and enlarge our training offerings after the installation of the systems"* (female, mid 30s). The current training strategy of the healthcare provider concentrates on the phases *pre-implementation, implementation and follow-up*. To cover all six phases of technology use for analyzing appropriate training methods of the healthcare professionals provides new insights for the management. In addition, the extent of data collection is desirable such that further factors are included in the methodology. Thus, the management asked to consider the willingness-to-pay in the survey. A manager said "*So we can see if this factor [willingness-to-pay] impacts the training and the appreciation"* (male, mid 40s). As the methodological approach appears as appropriate for the practical use, it can be adjusted and extended by further factors.

7 Discussion

By following the guidelines for design science research [14], a semi-structured, explorative survey methodology is designed. The designed methodology artifact provides an answer to the research question of how to analyze appropriate blended learning methods for technology-related training across the phases of technology use. We test the designed methodology for its practical application and evaluate the usefulness, comprehensibility and practicability. The resulting lessons learnt based on the first evaluation confirm the usefulness of the designed methodology for practice. The six phases of technology use are well accepted by the interviewees and provide guidance for the appropriateness of blended learning methods over a longer timeframe as well as the identification of technology-related training needs.

As the methodology concentrates on the analysis of how to deliver training and the appropriateness of online and traditional learning methods we contribute to TNA research. The findings of our literature review (see section 2.1) shows that the focus of TNA is mainly on what training to deliver such as knowledge and skills. In addition, the planning is often short-term oriented. Also in technology implementation research the training of new technology is mainly concentrating on the implementation phase. Therefore, we enlarge TNA research as well as technology implementation research by integrating further aspects such as the long-term orientation of training needs and the appropriateness of traditional and online learning methods across the phases of technology use. The concept of TNA provides the basis for analyzing training needs with regards to the appropriateness of learning methods. In addition, the six phases of technology use were integrated into the semi-structured, explorative survey methodology. The designed methodology addresses the research request of Santhanam et al. [3] regarding the match of training methods with individual needs and characteristics. Consequently, the results provide implications for the training design such that the appropriate amount of training is considered within the conceptualization of a training method. Further contextual factors such as characteristics of the target group (e.g. user groups, culture) [28-30] can be addressed in the training design to ensure positive learning outcomes. Thus, future research can apply the methodology to develop and validate theoretical models about the appropriateness of learning methods in different contexts. In summary, we contribute with our semi-structured explorative survey methodology as a basis for further investigations of factors influencing the relation between learning methods especially of blended learning and learning outcomes.

Next to the theoretical implications, the application of the semi-structured, explorative survey methodology results in implications for practice. The interactive methodology is easy to implement for practical use. Based on the positive feedback gathered by the interviewees from different countries as well as from the management of the healthcare technology company, the methodology facilitates to reflect about past training methods used for the implementation of new technology.

As the semi-structured, explorative survey methodology is applied for the first time within the study, there are limitations regarding the validity and reliability. In addition, a complete evaluation following the requirements of design science research is necessary in future research as within this paper only a first application of the designed method was done that focused on the acceptance of the methodology by learners. For approving these factors further studies are necessary. Additional impact factors such as the working environment can be examined in-depth. Furthermore, the usefulness of the methodology itself for other contexts should be investigated. This enables in addition, the comparison of the appropriateness of blended learning contexts based on different contexts, technologies and users. Therefore, in future research, the methodology should be applied by conducting further studies. Based on the lessons learnt and resulting adjustments the methodology should be investigated in other countries or healthcare technologies to ensure further validation. In addition, the methodology can be extended by further factors such as willingness-to-pay.

References

1. Denby, S.: The importance of training needs analysis. Industrial and Commercial Training 42, 147–150 (2010)
2. Anderson, G.: A Proactive Model for Training Needs Analysis. Jnl Euro Industrial Training 18, 23–28 (1994)
3. Santhanam, R., Yi, M., Sasidharan, S., Park, S.-H.: Toward an Integrative Understanding of Information Technology Training Research across Information Systems and Human-Computer Interaction: A Comprehensive Review. Transactions on Human-Computer Interaction 5, 134–156 (2013)
4. Aguinis, H., Kraiger, K.: Benefits of training and development for individuals and teams, organizations, and society. Annual review of psychology 60, 451–474 (2009)
5. Reed, J., Vakola, M.: What role can a training needs analysis play in organisational change? Journal of OrgChange Mgmt 19, 393–407 (2006)
6. Graham, C.R.: Blended Leaning Systems: Definition, Current Trends, and Future Directions. In: Bonk, C.J., Graham, C.R. (eds.) The handbook of blended learning. Global perspectives, local designs. Pfeiffer, San Francisco (2006)
7. Welsh, E.T., Wanberg, C.R., Brown, K.G., Simmering, M.J.: E-learning: emerging uses, empirical results and future directions. International Journal of Training and Development 7, 245–258 (2003)
8. Brown, J.: Training Needs Assessment: A Must for Developing an Effective Training Program. Public Personnel Management 31, 569–578 (2002)
9. Cooper, R.B., Zmud, R.W.: Information Technology Implementation Research: A Technological Diffusion Approach. Management Science 36, 123–139 (1990)
10. Schwarz, A., Chin, W.W., Hirschheim, R., Schwarz, C.: Toward a process-based view of information technology acceptance. Journal of Information Technology 29, 73–96 (2014)
11. Blanchard, P.N., Thacker, J.W., Pichai, A.: Effective training. Systems, strategies, and practices. Pearson Education, Harlow (2013)
12. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly 37, 337–355 (2013)
13. Cook, D.A., Levinson, A.J., Garside, S., Dupras, D.M., Erwin, P.J., Montori, V.M.: Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. Academic medicine : journal of the Association of American Medical Colleges 85, 909–922 (2010)
14. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28, 75–105 (2004)
15. Vaishnavi, V. and Kuechler, B.: Design Science Research in Information Systems, http://www.desrist.org/design-research-in-information-systems/ (Accessed: 29.10.2016)
16. Goldstein, I.L.: Training in organizations. Needs assessment, development, and evaluation. Brooks/Cole Pub. Co., Pacific Grove, Calif. (1993)
17. Byham, W.C., Adams, D., Kiggins, A.N.: Transfer of modeling training to the job. Personnel Psychology 29, 345–349 (1976)
18. Ford, J.K., Noe, R.A.: Self-assessed training needs: the effects of attitudes toward training, managerial level, and function. Personnel Psychology 40, 39–53 (1987)

19. Phang, C.W., Kankanhalli, A., Ang, C.: Investigating organizational learning in eGovernment projects: A multi-theoretic approach. The Journal of Strategic Information Systems 17, 99–123 (2008)
20. Waldman, D.A., Marino, F.J., Avolio, B.J.: A Multiple level of investigation of personnel ratings. Personnel Psychology 43, 811–835 (1990)
21. Moore, M.L., Dutton, P.: Training Needs Analysis: Review and Critique. Academy of Management Review 3, 532–545 (1978)
22. Clark, R.C., Mayer, R.E.: e-Learning and the Science of Instruction. Pfeiffer, San Francisco, CA, USA (2011)
23. Moreno-Ger, P., Torrente, J., Bustamante, J., Fernández-Galaz, C., Fernández-Manjón, B., Comas-Rengifo, M.D.: Application of a low-cost web-based simulation to improve students' practical skills in medical education. Int Jnl Med Infor 79, 459–467 (2010)
24. Melton, B.F., Bland, H., Chopak-Foss, J.: Achievement and Satisfaction in Blended Learning versus Traditional General Health Course Designs. ij-sotl 3 (2009)
25. Al-Qahtani, A.A., Higgins, S.E.: Effects of traditional, blended and e-learning on students' achievement in higher education. Jnl of Computer Assisted Learning 29, 220–234 (2013)
26. Kraiger, K., Ford, K., Salas, E.: Kraiger- Application of Cognition skill-based and affective theories of learning outcomes to new methods of training evaluation. Journal of Applied Psychology Monograph 78, 311–328 (1993)
27. Arbaugh, J.B., Godfrey, M.R., Johnson, M., Pollack, B.L., Niendorf, B., Wresch, W.: Research in online and blended learning in the business disciplines: Key findings and possible future directions. The Internet and Higher Education 12, 71–87 (2009)
28. Buyl, R., Nyssen, M.: MedSkills: a learning environment for evidence-based medical skills. Methods Inf Med 49, 390–395 (2010)
29. Lopez-Campos, G., Lopez-Alonso, V., Martin-Sanchez, F.: Training health professionals in bioinformatics. Experiences and lessons learned. Methods Inf Med 49, 299–304 (2010)
30. Sánchez-Mendiola, M., Martínez-Franco, A.I., Rosales-Vega, A., Villamar-Chulin, J., Gatica-Lara, F., García-Durán, R., Martínez-González, A.: Development and implementation of a biomedical informatics course for medical students: challenges of a large-scale blended-learning program. JAMIA 20, 381–387 (2013)
31. Sharma, R., Yetton, P.: The Contingent Effects of Training, Technical Complexity, and Task Interdependence on Successful Information Systems Implementation. MIS Quarterly 31, 219–238 (2007)
32. Alter, S.: Work System Theory: Overview of Core Concepts, Extensions, and Challenges for the Future. Journal of the Association for Information Systems (14:2), 72-121 (2013).
33. Bruque, S., Moyano, J., Eisenberg, J.: Individual Adaptation to IT-Induced Change: The Role of Social Networks. Journal of Management Information Systems 25, 177–206 (2008)
34. Li, X., Hsieh, J. J. P.-A., and Rai, A.: Motivational Differences Across Post-Acceptance Information System Usage Behaviors: An Investigation in the Business Intelligence Systems Context. Information Systems Research (24:3), 659–682 (2013)
35. Fichman, R.G., Kohli, R., Krishnan, R.: Editorial Overview —The Role of Information Systems in Healthcare: Current Research and Future Trends. Information Systems Research 22, 419–428 (2011)
36. Wilson, E.V., Lankton, N.K.: Interdisciplinary Research and Publication Opportunities in Information Systems and Health Care. CAIS 14, 332–343 (2004)
37. Lewis, R.J., Tamborini, R., Weber, R.: Testing a Dual-Process Model of Media Enjoyment and Appreciation. Journal of Communication 64, 397–416 (2014)

About Well-considered Decisions, Favorable Alternatives and Sudden Ideas: A Qualitative Research to Identify Beliefs that Influence Women to Study Information Systems in Germany

Caroline Oehlhorn, Sven Laumer, and Christian Maier

University of Bamberg, Germany
{caroline.oehlhorn,sven.laumer,christian.maier}@uni-bamberg.de

Abstract. Our research reveals motivational beliefs that influenced the decision to study information systems of female students in Germany. We apply a qualitative methodology in form of narrative interviews and interviewed 21 female students at a German university. On the basis of the theory of planned behavior, our results expose that a range of different beliefs exist which influenced the women's decision. We are able to expand the results of previous studies to the German context regarding especially normative beliefs, alternate choices or prior experiences dealing with information technology. In addition, we also uncover a couple of new beliefs that had not been revealed before.

Keywords: Women and IT, IS enrollments, study subject selection, theory of planned behavior.

1 Introduction

Recent articles coined the phrase *War for Internet Talent* [1] reflecting that the competition of well-educated information systems (IS) professionals reached a new level. Among others, recent statistics reveal that organizations cannot fill their information technology (IT) vacancies [2]. In line with that, top managers consider the education of IS professionals as one of the major issues [3]. As particularly women are underrepresented in the IT profession, institutions established plenty of women-specific programs, such as *Mädchen und Technik* (German for 'Girls and Technology') or *Girls Day*, to motivate women to study IS. Such programs were necessary as women and men base their decision on different aspects [4, 5].

From a theoretical perspective, the reasons why women focus a career within the field of IT are a well-discussed topic within IS research so far [6, 7]. Nevertheless, there exists also a wide range of reasons that seem to prevent women from doing so [8–13]. Respectively to the choice of studying IS, these include nonexistent interest, missing familiarity with IS or the perception of insufficient computer skills [10].

On closer consideration of the situation in Germany, we notice a slight increase regarding the number of enrollments in IS study programs of young women. Since

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Oehlhorn, C.; Laumer, S.; Maier, C. (2017): About Well-considered Decisions, Favorable Alternatives and Sudden Ideas: A Qualitative Research to Identify Beliefs that Influence Women to Study Information Systems in Germany, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 365-379

1998, the number of German women studying IS has been quintupled until today, whereas the percentage of women increased slowly from 12.8% to 17.8% [14].

A further increase in this context is desired by both, business and society. To sustainably support this process, it might not be enough to solely implement measures, but also investigating possible causes. As a first step into this direction, we want to detect the women's reasons that motivate them to matriculate into IS study programs. Only few research aims to identify motivational factors for studying IS, although previous research calls up to explore this topic in detail [15]. We, therefore, propose the following research question:

Which beliefs motivate women in Germany to study information systems?

In order to answer this research question properly, we conducted narrative interviews with 21 female students from a German university that offers two IS bachelor and three IS master degree programs. As the decision to study IS is usually planned and consciously made, our research orientates towards the theory of planned behavior (TPB). The structure of our paper is as follows: firstly, we explain the theoretical background focusing TPB and previous research related to IS studies. Next, we illustrate the applied methodology in section three and present our results in section four. The results are discussed in section five followed by a short conclusion.

2 Theoretical Background

As we want to elaborate the motivations that drive women to decide for studies in the field information systems today, the status quo of related existing literature will be presented first. The decision to study a certain subject is a planned and conscious decision [16–18]. Therefore, we refer to TPB as the theoretical foundation to explain the conscious decision studying IS. In the following subsections, we reveal theoretical essentials about TPB and decisions for academic studies among students seen from an IS-related perspective.

2.1 Theory of Planned Behavior

TPB is a frequently used theory within leading IS research [19, 20] that explains human behavior connecting beliefs and behavior. According to TPB, a person's behavior and previous intentions can be predicted by personal and environmental factors named attitude towards this certain behavior, subjective norm and perceived behavior control (PBC) [18]. TPB proposes that behavior is influenced by behavioral intention, which is influenced by the subjective norm regarding this behavior, the attitude toward this behavior, and perceived behavioral control (PBC) in turn. Subjective norm is related to social pressures a person is confronted with when deciding whether to perform a certain behavior or not. The attitude toward this behavior is defined as the person's propensity toward the behavior in relation to the expected consequences. PBC is seen as the person's perception of the difficulty of performing this behavior. TPB is useful to address our research question as it predicts intentions and behaviors very accurately [21].

In its expanded version that we apply to our research, the three influential factors within the TPB model are based on antecedent beliefs: behavioral beliefs affect the attitude, normative beliefs affect subjective norm, and control beliefs influence PBC [18]. Behavioral beliefs are individual beliefs of a person that address the related consequences of a certain behavior. Normative beliefs reflect beliefs of other people from the social environment that a person perceives regarding this particular behavior and control beliefs balance factors that influence the individual's feasibility of performing this behavior.

Referring to previous research that focuses on individual beliefs to explain behavioral choices concerning career issues [5, 10, 17, 22], we consider TPB as a proper theoretical base to examine our research question.

2.2 Previous Research

Previous research reveals a range of different factors that influence US students to study IS as major or to pursue an IT profession [23, 24]. Among others, certain factors show a significant impact on students' motivation to study IS: general interest, job-related and experiential beliefs, content-related and control beliefs, social impact and attitude in general [5, 8, 17, 23, 25–27]. Nonetheless, these results are limited in two ways. First, a comprehensive study that explains why women start studying IS does not exist yet as previous research only focusses particular beliefs. And second, previous research concentrated solely on US students [5, 10, 17]. However, these results cannot be applied for the decision of women who plan to study in Germany. While students in the US are more likely motivated to study out of reasons regarding wealth (to earn a high salary) and social status [28], students in Germany do so with regard to their perceived talents and personal development [29]. In addition, the educational systems of Germany and other European countries vary significantly from those in the US: students in the US start their academic studies attending some general and preparatory courses before choosing their major, students in Germany decide for their discipline of study before registering at a university in turn.

In this context, research also shows that these motivational factors differ by gender [4, 5]. Men, for example, are more motivated due to factors in the work itself. Women, by contrast, are more attracted by factors around the job such as job security or flexible working hours. As the demand for more women in the IT profession is still unabated, research also looks at the reasons why women are less attracted by the field of IS [8–13], suggests and analyzes countermeasures [12, 30]. As already mentioned, previous literature predominantly investigated the situation at numerous US universities, where the context differs due to the fact that the US university system distinguished from the German system. Furthermore, students in the US intend to study majors with regard to their expected wealth (to earn more money) and social status [28]. By contrast, students in Germany more likely focus on their perceived talents and personal development [29]. It is hitherto unresolved if the same factors that affect the decision towards study majors in the US are fully transferable for the decision of study subjects in Germany.

3 Methodology

We use a qualitative methodology by conducting interviews in the special form of narratives with 21 female students from a German university to reveal crucial beliefs based on TPB.

We chose narratives as a qualitative methodology for our research since narratives in their original form – a series of events in a specific order [31] – approximate the process that we aim to retrieve from the interviewees. The narrative approach[1] intends to conduct open-ended interviews that focus on the interviewees' previous experiences. As these experiences are mandatory to tell a story in the form of narratives, we focus on female students who have already decided to study IS. Some demographic information about the interviewees is depicted in Table 1. This procedure ensures that the interviewees possess the necessary experience to answer our interview questions. Furthermore, the narrative approach has been successfully applied to IS research to uncover and survey qualitative data at an early stage of research [32].

Initially, the call for participation was posted within the university's internal e-learning system and on the faculty's Facebook page in July 2016. Those students who answered the call for participation were invited to the chair's laboratory to take the interview. The interviews were recorded anonymously on tape with the agreement of each participant. According to the procedure of narratives, the interview started with some preliminary remarks that aim to encourage the participants to tell about their motivations, experiences, and thoughts before they decided to study IS. During their report, the interviewees were not interrupted to not interfere their process of self-reflection. When the interviewees finished, the interviewers asked additional questions to enlarge upon certain points if required.

All interviews were conducted during the month of July and lasted between 10 and 25 minutes. Afterwards, the interviews were transcribed verbatim and coded following qualitative coding procedures by using both descriptive and interpretative coding techniques [33].

In the first step, we searched for mentioned beliefs that obviously influenced the decision to study IS within the transcripts. For this purpose, we used a coding scheme that we adjusted to already proved beliefs of TPB obtained from thematically related literature [5, 10, 17]. Further lines and paragraphs, which did not match any existing beliefs withal, were separately coded and shown. Next, each of these newly identified beliefs was examined in its individual case. We searched for similarities and coherencies between those beliefs and grouped them into common categories referring to existing literature wherever possible. These categories were then associated with the three respective factors of TPB (behavioral, normative and control beliefs). To ensure an unambiguous assignment of the indicated beliefs and group categorization, coding was first done by all authors separately, and together at a later point in time.

[1] For more information about performing narrative approaches see [32].

Table 1. The interviewees' demographics

Participant	*Age*	*Study semester, intended degree*	*Participant*	*Age*	*Study semester, intended degree*
P 1	25	3, master	P 12	23	7, bachelor
P 2	26	4, master	P 13	25	7, bachelor
P 3	22	6, bachelor	P 14	22	1, bachelor
P 4	23	7, bachelor	P 15	22	1, bachelor
P 5	25	8, bachelor	P 16	19	1, bachelor
P 6	22	7, bachelor	P 17	24	7, bachelor
P 7	24	1, master	P 18	21	4, bachelor
P 8	24	2, master	P 19	23	8, bachelor
P 9	25	2, master	P 20	22	8, bachelor
P 10	21	4, bachelor	P 21	25	3, bachelor
P 11	21	1, bachelor			

4 Results

The outcomes of the interviews are shown in this section clustered by the different beliefs according to TPB (see Figure 2). As TPB is not tied to this issue, we examine this context aiming to reveal why women study IS (behavior). This is influenced by the intention to study IS which is affected by subjective norm, the attitude toward studying IS and the PBC toward studying IS in turn. By analyzing the conducted interviews, we want to point out which beliefs act as main drivers for these.
In each section, we give examples of statements that were narrated during the interviews. As we are not able to present all statements in this section, we included a comprehensive table (Table 2) in the appendix.

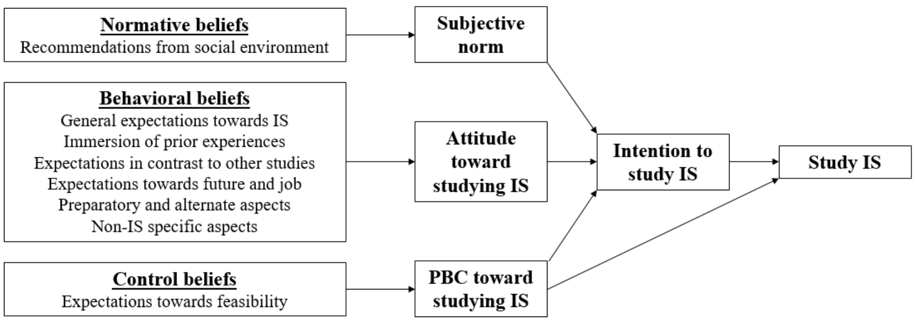


Figure 1. Results clustered by normative, behavioral and control beliefs based on TPB [18]

4.1 Normative Beliefs

As described in section 2.1, persons form normative beliefs based on the beliefs of other people from their social environment [18]. Following previous research [5, 17], we interpret those beliefs as normative beliefs when the influence of another person, who for example studies the same subject or gives recommendations, on the own behavior is recognizable. We especially notice normative beliefs in form of recommendations from the interviewees' social environment that called the students' attention towards IS or motivated them to enroll in IS studies. The mentioned social environment includes family members, friends, and persons in the university environment such as enrolled students, professors or members of the student council. Among all statements, we recognize that **recommendations from family members and friends** occurred predominantly.

As I said, I started with business administration, but this was some kind of temporary solution. Later, I met a friend at the university who studied information systems and she really raved about her studies with all the trimmings. So, I just went to one professor to get more information, and hereinafter, I decided to study the same. [P7]

I spoke to my brother who studies mechatronics where informatics is also included. He showed and explained me a lot and I was able to try some stuff. That was kind of funny. [P9]

4.2 Behavioral Beliefs

Behavioral beliefs are a person's individual beliefs that refer to the consequences of a particular behavior [18]. We expound those beliefs as behavioral beliefs when these are connected to aspects which can be seen as consequences of the decision for IS studies. The quantity of statements that belong to the category of behavioral beliefs was enormous and substantially widespread. We therefore clustered those into several subcategories that rely on different expectations or consequences related to an enrollment in IS study programs: expectations towards the study subject in general, in contrast to other study programs, or towards the own future and job possibilities as well as immersions of prior experiences with IT, IS studies as an alternative or preparatory study program, but also beliefs that cannot be related directly to the study subject.

General Expectations Towards IS. Previous research indicates that students fuel different expectations that can be related to the study subject itself such as the study contents or the composition of scientific fields [5, 10]. The interviewees within our research mentioned that they initially had different expectations on their enrollment in IS studies. They looked forward to the widespread range of topics, the composition of courses or the interdisciplinary field of the study subject for example. Others aimed to satisfy their genuine interest about IT, wished "to learn IT" or viewed the studies as a possibility of further education or even as an experiment.

Information systems is exciting related to its possible applications. I wanted to accept a new challenge after graduation from school and within information systems, I saw the best possibilities to improve myself. [P5]

Information systems offers a multifarious topic area, human-computer interaction, programming, media informatics for example. Hence, I thought, maybe I did not know all the different areas, and I wanted to try out if this is something for me. [P19]

I neither had economy at school nor informatics. But I was always a fan of: 'Let's do something new.' (...) I was really uneducated at the technical level as I never had a connection to it. But I was always interested. (...) I even had difficulties dropping subjects in school, and in this context, information systems is interesting as it combines two very different disciplines. [P8]

Immersion of Prior Experiences. Gaining first experiences, for example in form of introducing practices or even internships, in a certain subject has an evident influence on the career choice of young people [34]. Consequently, a couple of students already had first experiences with IT. They wanted to expand or deepen their knowledge with enrolling in IS studies. Other interviewees report that they had ease or fun learning IT at school.

I lived in the Silicon Valley as an au pair girl and had the perfect insight into this world and these themes in a way and I was deeply interested in this. Back home, I found this study program and it immediately appealed to me. [P21]

I actually had not that many expectations, because I did not know at all what is going to happen. All I knew, was that these two subjects (N.B.: Economy and computer science) are fun, and I hoped that it would still be fun. [P10]

I already had information systems and economy in school, and this was really fun to me. Information systems at school was more theory-loaded, I mean, we never did programming. We rather did modeling and this was quite cool. [P18]

Expectations in Contrast to Other Studies. Today, students are able to decide between a multitude of different study programs. Some programs are attended by plenty of students and can be described as almost overcrowded, whether other programs are allocated to niche areas with a manageable number of students. In this context, prior research shows findings revealing that students also decide for IS studies because they expect different circumstances from doing so in contrast to other study programs [5]. Within our research, the interviewees frequently mentioned statements that covered comparisons of IS with other study subjects such as business administration or informatics. Those comparisons referred to beliefs that other study programs are overrun with students or are not enough challenging (especially business administration) and in turn, some other studies would be too difficult for them to complete (especially informatics).

It's the whole direction of information systems, also how it is acknowledged within the economy and what I'm able to do with it. Because… When I think about this *bachelor in business administration, I have to say, I never liked it. Let us just put it in this way: this is just mainstream.* [P14]

Well, I first studied business administration (...). On the one hand, a lot of people study business administration. (...) IT is an issue where many people think it is all about programming and you have to do a lot of math and this is kind of scaring. But I think information systems shows that IT is so much more than just programming in the meantime. (…) *We just knew, that we would have more possibilities…and that enterprises would come up to us. And that the supervision is so much different with the result that I think I personally would be able to learn more.* [P13]

I ended up with information systems because I was searching for a study program that also included some IT components. But after I made some inquiries, I knew that computer science is really IT-loaded und you have to do lots of programming sometimes. Programming at school was not that much fun and therefore, I wanted something that also included some components of business administration. [P18]

Expectations Towards Future and Job. The choice of a study program is often a career decision at once. Students do not only choose their study subject for the next years, but set a path in the direction of career. Thus, students predominantly have different expectations towards their future career or certain jobs in mind when they think about enrolling in a study program [5, 17]. This corresponds to our results as the most often mentioned statements (see Table 2) of the interviewees refer to beliefs that are linked to their personal future. This shows that the students had already their lives after their studies in mind before they enrolled. Some already had specific perceptions how IS studies would affect their career or which (job) opportunities would be offered to them in the future.

You just have a positive feeling, if you can say: 'Well, cool, even if I go through with this, I'll get a job and because this is a thematically widespread study program, I would also be able to choose completely free what I am going to do.' [P11]

My cousin works as an IT business engineer and I just asked him what he is doing there. He just said in a platitudinous way: 'Well, I am sitting between a computer scientist and a business economist, and I explain to the computer scientist what the business economist wants to have.' That sounded pretty cool, especially because informatics was really easy for me to learn at school. [P1]

Preparatory and Alternate Aspects. Not all students are able to study their preferred study program from the start. Some programs require excellent grades at school or existing knowledge for example. And equally, some jobs are only allocated to candidates when they studied a certain subject in advance. So, students also consider study programs with respect to their existing and target-aimed opportunities [35]. In other cases, the interviewees told that they chose IS studies with regard to further purposes. Some students preferred other study programs and IS was just an alternative choice. For others, studying IS is a preparatory measure as it presents a mandatory qualification for specific master studies or desired careers that they already had in mind.

I initially planned to study medical engineering, but these programs only start in winter semesters. So, I examined possibilities at my hometown's university if there is something offered that would be preparatory for this. I found information systems and planned to do this for one semester with the opportunity to change later. [P16]

When studying IS, we have also the possibility to dabble in other areas and that was really interesting for me, because I was always interested in psychological studies. But my graduation grade was not good enough to enroll in psychology. So, this was also a nice opportunity to try psychology tentatively. [P19]

Since I was 15 years old, I know that I want to work in the gaming industry. There are people who do the programming, and there are people who do the management. And that is what I want to do. But to end up in this industry, you need a sort of special master, in gaming or something like that. And for doing a gaming master, you need a bachelor in computer science or just information systems. [P17]

Non-IS Specific Aspects. Whereas most beliefs can be associated with the study program itself, some interviewees also argued reasons that show no direct connection to IS studies. For example, some students were searching for study programs in a specific city or chose IS due to admission requirements. A couple of interviewees also reported that enrolling in IS was some kind of spontaneous or intuitive decision.

So at that time, I decided to study information systems kind of spontaneously. I always wanted to do something with business and first, I considered math or business mathematics. But it did not really appeal to me and I considered information systems in the short term instead. (…) I knew that I liked math und everyone told me, if I liked math, I should orient towards a technical direction and so, the decision resulted in information systems. [P12]

As I absolutely wanted to study in this city, and this study subject was a new offer at the university, I just applied there. In addition, the university over there sent off promises really early and that is why I did not apply elsewhere. [P1]

4.3 Control Beliefs

Fundamental research indicates control beliefs as factors that weigh a person's own feasibility of performing a behavior [18]. Prior research also understands the students' perceived feasibility of accomplishing a study subject as a control belief [17]. **Expectations towards the feasibility** of IS studies were often mentioned by the interviewees which is why we interpret those as control beliefs. Some students justified their choice based on good grades at school, existing knowledge or thorough information about the study program. In contrast to that, other interviewees believed that any study subject would be connected to obstacles anyway or they were going to learn everything during their studies and would not need any preparatory knowledge.

I really had no plan what to study. This sounds kind of haughtily, but I had good marks in all subjects at school. So, regarding my marks, it did not matter and it was not recognizable that anything was particularly suitable for me. [P6]

At that time, I completely reoriented and searched for mathematical study subjects, because math always suited more to me than languages, although I always had fun with that. (...) I thought that, in the worst case, I could not lose anything, because I also study business and I would have the opportunity to switch to economics or business administration. [P11]

I had no previous knowledge or anything. Except for one year at school, we did nothing in this area. Just some Word, but nothing 'IT-like'. That is why did not want to purely do computer science. Then, I found information systems that contains an imaginable scope of IT that would be possible for me. [P4]

When considering only the name, you can draw conclusions, if you are able to handle it or not. So, I informed a bit using the homepage and looked, if this would be feasible or if it would be lots of programming. [P2]

5 Discussion

Our research aim was to identify beliefs that motivate German women to start an IS study program. Existing literature shows that there is a wide range of different reasons why students-to-be choose IS or not, and how these differ by gender [5, 8, 17, 24]. However, these studies expose this phenomenon focusing only on US universities. As our research aims to uncover the beliefs of German students, we conducted narrative interviews with 21 female participants from a German university that offers several IS study programs. The completion of our research allows us to give a number of theoretical contributions and practical implications. Additionally, possible limitations of the current study are explained.

5.1 Theoretical contributions

First, we identified several beliefs that led to an enrollment in IS studies by interviewing 21 female students. After identification we defined each belief and grouped them into preceded categories with respect to previous literature (see Table 2). We integrated these categories in the three comprehensive beliefs according to TPB [18] referring to existing research within this issue [5, 10, 17, 23, 25, 27]. Our results show that the interviewees decided for IS studies based on recommendations from the social environment (normative beliefs), their perceived feasibility of the subject (control beliefs) and several behavioral beliefs such expectations towards the subject itself, towards future and job opportunities and in contrast to other study programs or immersions of prior experiences.

Second, we also contribute by confirming that the consideration of beliefs is a proper theoretical base for revealing intentions and behavior of women towards studying IS as previous literature has already shown before [5, 8]. To dissociate from those, which address solely the US university system, we conducted our research at a German university to examine the situation in Germany. Our results correspond with a couple of the already validated beliefs from previous research [5], such as job availability, aptitude, genuine interest and recommendations from the social environment. But we also see categorical differences such as formed normative beliefs out of recommendations, the students' personal and social image being an IS student [5] or negative beliefs concerning their performance in form of anxiety or low computer self-efficacy [8]. As a result, we also obtain the insight that cultural conditions might influence the women's decision to study IS which should be investigated in future research.

Third, we disclosed some other factors that influenced the interviewees' decision, which had not been considered by previous literature in this context before [5, 8, 12, 13, 17]. One is the existence of prior experiences in the field of IT: some students mentioned that they already had IT subjects at school or other impressions, such as internships or insights in IT enterprises, and that these were determinative for their decision for IS studies. Furthermore, the pure interest in the field of IS was part of many students´ decision making. Hereby, they mentioned that they had to exert their selves no matter which study programs they would choose or that they would be able to change their subject if they do not like it as already achieved grades could also be credited in other study programs such as business administration. Hereby, a certain serenity becomes apparent that further motivates the students to try IS studies, although some concerns regarding this behavior might exist.

5.2 Practical implications

We can also provide practical implications for universities, employers and other institutions that aim to attract women for an IT career. Our study results unveil that many interviewees perceived an IS study program as feasible after they sought adequate information. Several sources for information were mentioned including statements of IS students, online reports, brochures and others. As a consequence, institutions need to highlight possible job opportunities within the IT field, different fields of application and other benefits that would be offered to women if they decide to study IS.

On a final note, we would like to make all girls, who are interested in an IS study program, aware that none of the interviewees did regret her decision yet, even if there had been little concerns in advance. By reciting one student, who said *'Actually, it was a gut decision. I did not know for what I let myself in but in the end… well, it just seems to be the right thing'* [P6], we would like to note that a certain (and wise) degree of serenity is not wrong when facing important decisions, especially if these affect the own future. Sometimes it is just about trying something new or succumb to the own interest.

5.3 Limitations

The research on hand is subject to several limitations. Firstly, our study was only conducted at one university in Germany as our research's focus is on the choice of the study subject, not the study location. Nonetheless, we cannot ensure that our results are completely transferable to other universities. We also abandoned to interview participants who decided against an IS study program due to the choice of our methodological approach in form of narratives as it is explained in detail in the third section. In addition, we interviewed only female students as this is done by leading research focusing women issues within the IS field [7, 36].

6 Conclusion

We aimed to answer the research question: "Which beliefs motivate female students in Germany to study information systems?" with our research by applying a qualitative methodology in form of narrative interviews with 21 female students. Our results show that a range of different beliefs influenced the women's decision to study IS. Those exhibit several similarities and differences toward previous studies that have already been conducted at universities in the US.

References

1. Efrati, A. and Tam, P.-W.: Google battles to keep talent, http://www.wsj.com/articles/SB10001424052748704804504575606871487743724
2. van Heur, R.: Fears of software skills shortage in Germany and the Netherlands, http://www.computerweekly.com/news/4500269840/Fears-of-software-skills-shortage-in-Germany-and-the-Netherlands
3. Luftman, J.N., Kempaiah, R.M., Rigoni, E.H.: Key issues for IT executives 2008. MIS Quarterly Executive 8 (2009)
4. McKinney, V.R., Wilson, D.D., Brooks, N., O'Leary-Kelly, A., Hardgrave, B.: Women and men in the IT profession. Commun. ACM 51, 81–84 (2008)

5. Zhang, W.: Why IS: Understanding undergraduate students' intentions to choose an information systems major. Journal of Information Systems Education 18, 447–458 (2007)
6. Panteli, N.: A community of practice view of intervention programmes. The case of women returning to IT. Information Systems Journal 22, 391–405 (2012)
7. Quesenberry, J.L., Trauth, E.M.: The (dis)placement of women in the IT workforce. An investigation of individual career values and organisational interventions. Information Systems Journal 22, 457–473 (2012)
8. Roach, D., McGaughey, R.E., Downey, J.P.: Gender within the IT major - a retrospective study of factors that lead students to select an IT major. IJBIS 7, 149–165 (2011)
9. Ahuja, M.K.: Women in the information technology profession. A literature review, synthesis and research agenda. Eur J Inf Syst 11, 20–34 (2002)
10. Croasdell, D., McLeod, A., Simkin, M.G.: Why don't more women major in information systems? Info Technology & People 24, 158–183 (2011)
11. Trauth, E.M., Quesenberry, J.L., Huang, H.: A multicultural analysis of factors influencing career choice for women in the information technology workforce. Journal of Global Information Management 16, 1–23 (2008)
12. Clayton, K., Beekhuyzen, J., Nielsen, S.: Now I know what ICT can do for me! ISJ (Information Systems Journal) 22, 375–390 (2012)
13. Burke, R.J., Michie, S., Nelson, D.L.: Barriers women face in information technology careers. Women in Management Review 21, 10–27 (2006)
14. Statistisches Bundesamt: Studierende nach Nationalität, Geschlecht und Studienfach - Wirtschaftsinformatik, https://www-genesis.destatis.de/genesis/online/logon?sequenz=tabelleErgebnis&selectionname=21311-0003
15. Panko, R.R.: IT employment prospects. Beyond the dotcom bubble. Eur J Inf Syst 17, 182–197 (2008)
16. Crawley, F.E., Black, C.B.: Causal modeling of secondary science students' intentions to enroll in physics. J. Res. Sci. Teach. 29, 585–599 (1992)
17. Heinze, N., Hu, Q.: Why college undergraduates choose IT. A multi-theoretical perspective. Eur J Inf Syst 18, 462–475 (2009)
18. Ajzen, I.: The theory of planned behavior. Organizational Behavior and Human Decision Processes 50, 179–211 (1991)
19. Mathieson, K.: Predicting user intentions. Comparing the technology acceptance model with the theory of planned behavior. Information Systems Research 2, 173–191 (1991)
20. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: Toward a unified view. MIS Quartlery 27, 425–478 (2003)
21. Ajzen, I., Madden, T.J.: Prediction of goal-directed behavior. Attitudes, intentions, and perceived behavioral control. Journal of Experimental Social Psychology 22, 453–474 (1986)

22. Arnold, J., Loan-Clarke, J., Coombs, C., Wilkinson, A., Park, J., Preston, D.: How well can the theory of planned behavior account for occupational intentions? Journal of Vocational Behavior 69, 374–390 (2006)
23. Li, L., Zhang, C., Zheng, G.: Promoting information systems major to undergraduate students - A comprehensive investigation. Journal of Information Systems Education 25, 211–219 (2014)
24. Walstrom, K.A., Schambach, T.P., Jones, K.T., Crampton, W.J.: Why are students not majoring in information systems? Journal of Information Systems Education 19, 43–54 (2008)
25. Rouibah, K.: Understanding student drivers and obstacles toward MIS major from the perspective of an arab country: The case of Kuwait. Issues in Information System 13, 58–71 (2012)
26. Kuechler, W.L., McLeod, A., Simkin, M.G.: Why Don't More Students Major in IS? Decision Sciences Journal of Innovative Education 7, 463–488 (2009)
27. Ferratt, T.W., Hall, S.R., Prasad, J., Wynn, Jr., Donald: Choosing management information systems as a major: Understanding the smiFactors for MIS. Communications of the Association for Information Systems 27 (2010)
28. Astin, A.W.: The changing American college student: Thirty-year trends, 1966–1996. The Review of Higher Education 21, 115–135 (1998)
29. Hachmeister, C.-D., Harde, M.E., Langer, M.F.: Einflussfaktoren der Studienentscheidung. Eine empirische Studie von CHE und Einstieg. CHE, Gütersloh (2007)
30. Adya, M., Kaiser, K.M.: Early determinants of women in the IT workforce. A model of girls' career choices. Info Technology & People 18, 230–259 (2005)
31. Bennett, A., Royle, N.: An introduction to literature, criticism and theory. Prentice Hall Europe, London (1999)
32. Schwarz, A., Chin, W.W., Hirschheim, R., Schwarz, C.: Toward a process-based view of information technology acceptance. J Inf Technol 29, 73–96 (2014)
33. Myers, M.D.: Qualitative research in business and management. SAGE Publications Ltd, London (2013)
34. Carr, J.C., Sequeira, J.M.: Prior family business exposure as intergenerational influence and entrepreneurial intent. A Theory of Planned Behavior approach. Journal of Business Research 60, 1090–1098 (2007)
35. Young, M.R.: Choice-based segmentation as an enrollment management tool. Journal of Marketing for Higher Education 12, 69–83 (2003)
36. Trauth, E.M., Quesenberry, J.L., Huang, H.: Retaining women in the U.S. IT workforce. Theorizing the influence of organizational factors. Eur J Inf Syst 18, 476–497 (2009)

Appendix

Table 2. Statement overview

Statement	*Interviewee #P*	*Total*
Normative beliefs: *Recommendations from…[5, 17]*		**10**
…family members	1, 9, 11, 17, 19	5
…friends	1, 3, 5, 7, 11, 13	6
…university members	7, 8	2
Behavioral beliefs		**21**
Expectations towards IS in general[5, 10]		*13*
Interdisciplinarity, mixture of courses, widespread topic area	4, 5, 8, 10, 11, 17, 18, 19	8
Satisfaction of genuine interest	2, 5, 8, 10, 17	5
Possibility of personal development	5, 8	2
"Learning IT", insight into IT	4, 7, 8, 15, 17, 18	6
Experimentation, try something "new"	8, 14, 19, 20	4
Immersion of prior experiences[34]		*10*
IT at school	1, 5, 10, 17, 18	5
Further experience of the IT world	8, 21	2
Deepen existing knowledge	8, 19	2
Expand/transfer existing knowledge to the IT side	2, 15	2
Expectations towards IS in contrast to other study programs[5]		*8*
Versatility, better opportunities	2, 9, 11, 13, 14, 20	6
Higher challenging character	7, 9, 13	6
No overrun	13, 14, 15	6
Expectations towards future and job[5, 17]		*10*
Better/more interesting job opportunities	1, 3, 5, 11, 13, 15, 19, 21	8
Opportunity to help forming the future	2, 13	2
IT is necessary in every job	13, 15, 18	3
Preparatory and alternate aspects[35]		*6*
Preparation for specific master programs/job positions	6, 14, 15, 17	4
Alternative choice	14, 16, 19	3
Non-IS specific aspects		*9*
Admission and location specific aspects	1, 2, 14, 16, 18, 20	6
Spontaneous, intuitive decision	6, 8, 12, 14, 16	5
Control beliefs: *Expectations towards feasibility based on…[17]*		**19**
…low level of technical contents	4, 7, 11, 13, 18	5
…no required previous knowledge (IT is taught in courses)	3, 4, 14, 15, 19	5
…statements of others	3, 7	2
…beliefs that women are capable to do it	1, 3, 17	3
…beliefs that obstacles will occur in every study program	8, 19	2
…extensive information about IS studies	2, 5, 15, 18	4
…opportunities to change the subject/credit previous grades	5, 7, 9, 11, 13, 14, 20	7
…good grades at school	6, 17	2

When Life Gives You Lemons: How rating scales affect user activity and frustration in collaborative evaluation processes

Thomas Wagenknecht[1], Jan Crommelinck[1], Timm Teubner[2], and Christof Weinhardt[2]

[1] FZI Research Center for Information Technology, Berlin, Germany
{wagenknecht,crommelinck}@fzi.de
[2] Karlsruhe Institute of Technology, Karlsruhe, Germany
{timm.teubner,weinhardt}@kit.edu

Abstract. Initiators of open innovation processes involving customers or employees often face vast amounts of idea proposals. These proposals vary greatly in terms of quality, which is why organizers often engage the users themselves in the evaluation process. Building on the concept of information overload, we evaluate the effects of three distinct rating scales on users' activity and frustration measures. On the basis of an open innovation campaign for employees of a public-private institution in Germany, we systematically compare the novel "bag of lemons" method with conventional Likert scales and up-down-voting schemes. Our results demonstrate that the "bag of lemons"-approach yields higher levels of user activity, but is also perceived as significantly more frustrating. We find this effect to be fully mediated by perceived information overload, which points to potential avenues for the design of stimulating yet tolerably complex Information Systems for open innovation and rating techniques.

Keywords open innovation, rating scales, information overload, participation

1 Introduction

From strategic planning to product innovation, small and large firms as well as other organizations are involving their employees and stakeholders to propose novel ideas through digital platforms [1–3]. These processes are sometimes strictly limited to participation within the company or part of a larger open innovation campaign, including customers, suppliers, and other interested parties [4, 5]. Regardless of their target group, these platforms all have in common that users face vast amounts of proposals of varying quality, but only a few can or even should be implemented [2, 6]. Hence, there is a strong need for group decision support systems (GDSS) that enable users to filter ideas appropriately [7], i.e., that achieve high accuracy in identifying the best ideas and avoid to expose users to the adverse effects of information overload, including frustration and disengagement [8, 9].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Wagenknecht, T.; Crommelinck, J.; Teubner, T.; Weinhardt, C. (2017): When Life Gives You Lemons: How rating scales affect user activity and frustration in collaborative evaluation processes, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 380-394

Accordingly, there exists a myriad of filtering techniques. On the one hand, these include complex approaches such as prediction markets [10, 11], or automated methods like text mining that initially require a lot of human oversight and implementation capacity [12]. On the other hand, approaches like voting and user ratings are easier to implement and widespread on various online platforms. For instance, many social media and community platforms offer simple up- and down-voting (e.g., Reddit, Quora, Stackoverflow, etc.) or up-voting only (e.g., Facebook and Yammer in the form of "Likes"). Other platforms use Likert scales, often in form of star-ratings (e.g., Amazon, Airbnb, etc.). Yet, these methods face inherent shortcomings, including biased distributions [13], limited accuracy due to oversimplification, a possible disconnect between the goals of process organizers and raters, as well as reduced user satisfaction [14–16]. In this vein, the video platform YouTube dropped its Likert scale rating system in 2009 as users mostly rated content as either very bad or very good – rarely using any measures in the middle of the 5-point scale. Since then, the platform switched to up- and down-voting [17].

Seeking to address some of the shortcomings of existing approaches, Klein and Garcia [7] proposed a novel method. Their so-called "bag of lemons" (BOL) approach lets users in evaluation tasks allocate a predefined amount of *lemons* to those ideas they consider to be the worst. A lemon thus represents a negative assessment and a user can allocate multiple, indeed up to all of her lemons to one single idea. This way, the crowd is assumed to flag bad ideas, supposedly identifying a (remaining) set of high quality ideas. In fact, the BOL method outperformed Likert scales in terms of time for task completion and accuracy [7]. To follow up on these first auspicious insights, this paper systematically assesses the characteristics of the BOL method in comparison to up-/down voting and (conventional) Likert scales. In doing so, we focus on two factors. First, as crowd-based schemes rely on the laws of large numbers and the quality of collaborative evaluations usually increases in the number of independent assessments [18], we consider *user activity* under the three mentioned rating method regimes. Second, as crowd-based approaches typically work on a voluntary basis and hence require a positive user attitude and engagement [10, 19, 20], we consider the – potentially detrimental – effects on *frustration* as a key indicator of a non-positive attitude and user disengagement [20]. Such motivational variables are widely perceived as a crucial factor for user acceptance and usage of information systems [21, 22]. In this sense, this research is motivated by the following key drivers: First of all, there exists a clear research gap as BOL represents a novel method and its role in contrast to established methods is still unclear. However, organizations increasingly seek to involve their employees, citizens, or members in decision making in order to increase content, loyalty, identification, and productivity – often using those very collaborative voting techniques [23]. In consequence, as accruing informational charges grow constantly, such methods may expose participants to excessive informational load, yielding undesired results such as frustration, disaffection, and disengagement [8].

To connect the different rating methods with our target variables, we hence base our research on two intermediate, explanatory factors. First, as the BOL method represents a novel and commonly unknown rating technique, we consider the factor of *perceived novelty*, capturing potential user deterrence by the unknown, or a lack of comprehensibility. Second, as BOL requires users to deal with a host of informational bits and pieces, *information overload* may be a concern. It was shown to yield adverse effects on employees as they are exposed to ever-growing amounts of unrestricted and unfiltered data [8, 9]. Accordingly, in this study, we pose the following overarching research questions:

RQ_1: How does the "Bag of Lemons" rating method affect user activity and frustration in a collaborative evaluation task?

RQ_2: Which role do perceived novelty and information overload play in mediating these effects?

To address our research questions, we conduct an online-based field experiment, including the collection of survey data. As part of a real world open innovation campaign, employees of a private-public institution rated the idea proposals of their peers. We systematically vary rating method, using up-/down voting, Likert scales [24] and the BOL method [7]. We investigate the ramifications for user activity, frustration [20], and task completion time, taking into account the factors perceived novelty and information overload [8]. Exceeding previous studies [7, 15, 20], users in this scenario were not forced to rate all ideas, which promises a more realistic situation and novel findings. In consequence, this study makes three main contributions to the Information Systems (IS) literature. First, we evaluate a novel, thus hardly researched method of idea evaluation (BOL) in comparison to more established methods (Likert scales, up-/down voting) in terms of the important indicators *user activity* and *frustration*, which has not or only scarcely been assessed by extant literature. By integrating these opposing factors within a joint research model, we enhance the understanding of collaborative evaluation processes in view of differentiated rating regimes [5, 25, 26]. Second, by relating these key indicators to mediating factors, we provide starting points for understanding *how* the different rating methods affect the users' perceptions and behaviors. In particular, we identify perceived information overload as a potential mediating factor at play. Third, our study provides a show case of employee-driven innovation [27] and computer-supported organizational participation [28].

This paper is organized as follows. We outline related work and the theoretical background in Section 2. Section 3 then illustrates our study design. Section 4 presents the results of our study. Lastly, we discuss our findings in view of theoretical and practical implications, limitations, and starting points for future research in Section 5.

2 Theoretical Background and Related Work

In recent years, the IS literature has begun to systematically evaluate ways to exploit the wisdom of the crowd, including a broad strand of research on open innovation processes [5, 29]. Notably, a number of studies analyzed voting and rating techniques on open innovation contests [7, 10, 15, 20, 30]. Such approaches relate to GDSS in the sense that groups evaluate proposals which were generated by the group itself, which can have important ramifications due to personal or social attachment, preoccupation, and other biases [31, 32]. With the emergence of large-scale open innovation contests, IS research revived its investigation of rating scales. Several studies in this line of research evaluated both quality and task completion time with regard to different rating techniques [15, 20, 30, 33, 34]. In this section, we describe the theoretical background of the concepts and factors that form the basis of our study. We begin with a brief introduction of open innovation contests.

2.1 Open Innovation

Adamczyk et al. [5] define open innovation contests as IT-based and time-limited competitions by individuals or organizations calling on the general public or a specific target group to propose innovative solutions. Thereby the organizers make use of the expertise, skills, and creativity of distributed crowds. Engaging employees and customers in open innovation processes can have several benefits for the organizers, including increased loyalty, brand image, and success in recruitment [35]. For an open innovation contest to be successful, previous research identified a number of factors. Organizers, for instance, need to express a sense of urgency and establish a trusted environment [14, 36]. Moreover, users might be motivated by gaining access to the knowledge of experts as well as receiving appreciation for their input by peers and organizers of the process [37]. Furthermore, extant research also established that collaborative tools drive increase the quality of results in open innovation engagements [23].

Recently, several leading IT corporations engaged both their customers and employees in open innovation contests. For instance, IBM’s “Innovation Jam” resulted in 46,000 product ideas proposed by 150,000 participants [1], while users in Dell’s ongoing “IdeaStorm” have generated more than 20,000 suggestions for product improvements thus far [6]. Open innovation contests among employees of a company are one application of employee-driven innovation [27, 28]. In the broader context of computer-supported organizational participation, these contests can be a way to actively provide employees the means to be part of the decision-making processes of their workplace, which was found to be related to increased employee commitment and productivity [28].

Considering the vast amount of ideas, it becomes more likely that an open innovation contest will produce more superior solutions than an innovation process limited to only few innovators [38]. Thus, in line with the “wisdom of the crowds” paradigm, some user-generated ideas are able to compete with expert or core inside innovators [15, 25, 39]. However, assessing these crowd proposals can be costly. Robinson and Schroeder [40] estimate that large corporations take about four hours working time and $500 just to evaluate one idea. Yet, only few ideas are really worth increased attention. Prior research established that open innovation processes tend to produce large idea collections that are highly redundant and greatly vary in terms of quality [2, 20, 33, 39], where only 10-30% of the ideas tend to be of good or high quality [33]. Put figuratively, large-scale open innovation processes create excellent needles. They do, however, also create the corresponding haystacks. The main challenge then is to identify the valuable propositions. One common solution to this problem is to engage users in the evaluation process using voting and rating techniques [7, 10, 15, 20, 30, 34].

2.2 Rating Scales, Attitudes, and Intrinsic Motivation

The usage of rating scales transforms the process of idea evaluation into a concrete task of judgment, where individuals consider a finite set of alternatives [10]. In effect, this enables the organizers of open innovation contests to reduce their costs for idea evaluation by basing decisions on aggregated user ratings.

However, the gathered data may depend on the specific rating scale. Prior research suggests that rating scales are prone to selection biases and other dysfunctionalities [7, 10, 15, 20, 30, 33, 34]. For instance, some researchers claim that rating scales often fail to properly distinguish between medium/good and excellent ideas [7, 34]. Moreover, there may occur discrepancies between the initiator's and the participants' goals and intentions. While initiators would like the participants to evaluate as many ideas as possible thoroughly, the latter are restricted both in terms of time and information available to them. Hence, organizers need to take potential factors such as non-interest, distractions, lack of knowledge, and workload into account [7, 20]. In consequence, they need to communicate clearly what, why, and how they would like their participants to do specifically.

Nonetheless, evaluation tasks are often described poorly and hence remain fuzzy. The rating scale itself hence become an important factor as participants are searching for potential cues [41]. In fact, participants tend to develop attitudes toward rating scales based on characteristics such as graphical elements and input variables [10, 19, 20]. Attitudes, in turn, can affect cognition and behavior [42]. In this context, Riedl et al. [20] found that users perceive different rating scales as more or less exciting, entertaining, satisfying, and positive, which can be explained by flow theory [43], suggesting that people can become very immersed by an activity, accompanied by high concentration on a task, while losing self-consciousness. Koufaris [44] suggested that flow states are related to increased intrinsic enjoyment and perceived control. Both constructs are also related to intrinsic motivation [45]. IS research established intrinsic motivation to be an important factor in creating favorable user perceptions, intention, and actual system use [21, 22]. In contrast, all too simple or overwhelmingly complex systems may deter users from entering such states, rendering system use a frustrating experience which is in consequence unlikely to be continued. Several potential antecedents of frustration come to mind. Given the structure of evaluation tasks with many diverse ideas, information overload is a concern which we further outline in the next paragraphs.

2.3 Information Overload

Information overload can be characterized as a state in which cognitive processing capacity is exceeded by the volume and speed of incoming stimuli that need to be processed [8]. People continuously evaluate their usage of information systems and discontinue usage when experiencing techno stress [46]. For instance, Maier et al. [47] found that users stop using social network services when experiencing, among other factors, exhausting levels of information disclosures by friends leading to information overload. Koroleva et al. [48] found similar results for Facebook and Eckhardt el. [49]

did so, asking participants in an experiment on LinkedIn to extract specific information for a job application. The phenomenon of information overload might be especially pronounced in open innovation evaluation tasks as users need to process a manifold, diverse, partly contradicting, and often novel set of ideas. Aggravatingly, the proposers usually do not follow a common schema, style, or language in describing their ideas. Comparing ideas across one another may hence be particular challenging.

Depending on the structure of the rating scale and evaluation task, perceived information load may thus differ [8]. It has, however, not been investigated with regard to rating scales in IS studies thus far. In the following, we hence describe a design allowing to relate users' perceptions of information overload to different rating methods, forming the basis of the field experiment reported in this paper.

3 Experimental Design

In this section, we outline an approach to address our research questions. Similar to Klein and Garcia [7], our study is based on an (internal) open innovation campaign at an actual private-public research center. Both the ideation as well as the evaluation phase were part of a broader participatory process at this institution [28]. The institution is legally incorporated as a foundation, disposes over a yearly budget of approximately €14 million, and employs a total of 280 people. Employees work on a variety of projects in the domains of computer science, information technology, robotics, and engineering.

Our study employs a two-staged approach. In the first stage, employees of this institution were invited to propose ideas on how to make the research center an (even better) employer via an online system. We invited all employees to this online platform. In the second stage, all employees were invited again to rate the ideas in a condensed set, using either BOL, up- and down-voting, or Likert scales.

3.1 Stage 1: Idea Generation

Employees of the institution were asked to propose ideas on how to make the center an (even better) employer. In a first phase, we received a total of 71 "raw" proposals. Before proceeding to the second stage, we eliminated hoax and proposals not compliant with the terms of use (e.g., including clear names of employees or foul language), consolidated redundant proposals, redacted grammatical and other language- and style-related issues, and in consequence, generated a condensed and workable idea corpus of 42 proposals. The proposals covered a wide range of topics, addressing organizational procedures, marketing, human resources, and many other areas. In this first stage, participants were able to propose ideas within a range of two weeks. Ideas were generally posted anonymously in order to both comply with German data protection legislation and to enable employees to speak their mind freely [50, 51].

3.2 Stage 2: Idea Evaluation

In the second stage, employees were then invited to rate their peers' proposals on another online platform. This platform was accessible for two weeks, too. Here, each employee could participate only once. Participants were prompted to assess the ideas' overall quality, which may be based on subcategories such novelty, feasibility, or value to the company [15, 20, 30, 39]. Note, however, that these sub-dimensions were not surveyed separately. In fact, idea evaluation was based on either bag of lemons, up- and down voting, or Likert scales.

Each participant was allocated to only one of the three treatment conditions (between-subjects design). All participants were presented the same 42 proposals in all treatment conditions, using a random order for each participant in order to rule out sequence effects. Following Klein and Garcia [7], participants in the BOL setting disposed over a total of eight lemons, representing ~20% of the total idea basket, which they were able to allocate to the ideas. In the up- and down-voting setting, participants could either up-vote or down-vote each idea once. This setting replicates that of platforms such as YouTube. Participants in the Likert scale setting were able to rate the ideas on 5-point Likert scales, ranging from 1 (very bad) to 5 (very good). Exceeding previous studies [7, 15, 20], participants in the Likert and up- and down-voting treatments were free to rate as many ideas as they liked, that is, there was neither a minimum nor maximum requirement. Participants were asked to complete a mandatory quiz before the actual rating task in order to ensure comprehension and hence validity.

3.3 Measures

After completing the rating process, participants were asked to conduct a brief survey. To ensure validity, previously validated scales were used and adapted to the context of this study. We assessed user attitudes towards the rating method, operationalized by the categories novelty and frustration [20, 52]. Information overload was adopted based on the items proposed by Schultz and Vandenbosch [8]. To assess user activity, we measured how many votes were casted in relation to the maximum number of votes in the respective treatment. This index ranges between 0 and 1.

Table 1. Measurement items

Construct	***Item***	***Source***
Perceived Novelty	Using the rating scale was a novel experience to me.	[20]
Frustration	Using the rating scale was a frustrating experience to me.	[20]
Information Overload	In using the rating scale, I was forced to concern myself my many idea proposals.	[8]
	In using the rating scale, I could not focus on the actual relevant idea proposals.	
	The rating scale overcharged me by too many idea proposals and too much information.	

4 Results

In total, 141 participants completed the questionnaire, representing approximately 50% of the total workforce at the institution. Altogether, 54 participants evaluated the ideas using BOL, 48 were in the Likert treatment, and 39 in the up- and down-voting treatment. In compliance with German privacy regulation, participants were able provide personal information on a voluntary basis. Thus, only part of our sample reported age (61.5%) and/or gender (71.5%). The age of the (reporting) participants ranged from 18 to 37 years (mean 28.9). Moreover, 80% of our participants were male. These characteristics did not differ significantly among the three treatments.

We first turn to the central target measures of this study, user activity and frustration. As illustrated in Figure 1, user activity was highest for the BOL method, and lowest for up-/down voting. A set of t-test confirms the significance of these differences ($t_{BOL/Likert} = 1.648$, $p=.103$; $t_{BOL/U\&D} = 4.347$, $p<.001$; $t_{Likert/U\&D} = 3.206$, $p<.001$). As a first result, we thus note that the bag of lemons rating scheme facilitates higher levels of user activity than Likert scales or up- and down voting.

Next, we consider how frustrating users perceived the different rating methods. Figure 1 shows that BOL provokes markedly higher levels of frustration than the other methods, whereas Likert and up-/down voting yield comparable levels. A set of t-test confirms this impression statistically ($t_{BOL/Likert} = 2.498$, $p=.014$; $t_{BOL/U\&D} = 2.783$, $p=.007$; $t_{Likert/U\&D} = .283$, $p=.778$). As a second result, we note that the bag of lemons rating scheme facilitates higher levels of perceived frustration than Likert scales or up- and down voting.

Besides these focal measures, we surveyed the participants in terms of how novel and how (informational) overloading they perceived the three rating methods. As can be seen in Figure 1, both for novelty and information overload, the bag of lemons method yields (marginally) significant higher levels than the other two (Novelty: $t_{BOL/Likert} = 11.033$, $p<.001$; $t_{BOL/U\&D} = 11.711$, $p<.001$; $t_{Likert/U\&D} = .983$, $p=.328$; Overload: $t_{BOL/Likert} = 1.816$, $p=.072$; $t_{BOL/U\&D} = 2.555$, $p=.013$; $t_{Likert/U\&D} = 1.0613$, $p=.292$).

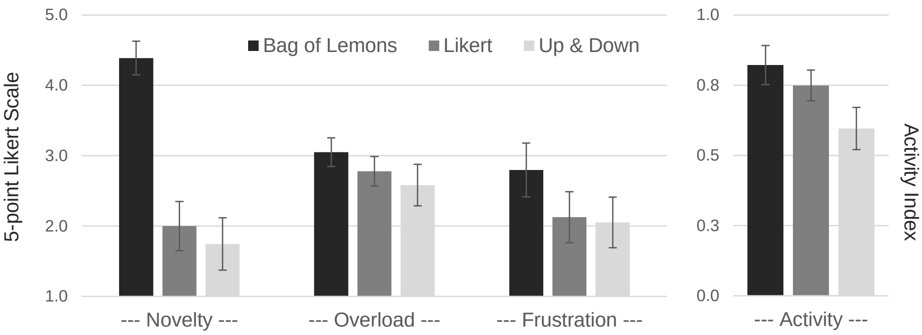


Figure 1. Overview of novelty, information overload, frustration, and activity scores (error bars indicate 95% confidence intervals)

We now turn to a structural analysis of the effects of rating scale on user activity and frustration. As we have outlined in Section 2, we hypothesize perceived novelty and information overload as potential mediators, that is, carriers and hence psychological determinants of the rating scale effects on the target measures. For doing so, we slightly simplify the analysis, comparing the bag of lemons method against both other methods simultaneously, that is, using only one binary dummy variable for "bag of lemons." Our model, along with the results, is depicted in Figure 2. We use structural equation modelling based on partial least squares (SEM-PLS) to operationalize this analysis. Specifically, SmartPLS 3.0 [53] was used due to its flexibility in terms of sample size and its lack of assumptions regarding data and residuals distribution [54]. The sample size of this study (n = 141) exceeded the minimum required to validate a model in PLS, given the present structural model [55]. Confirming the results from above, this analysis shows that the bag of lemons significantly increases the perception both of (rating scale) novelty (*b*=.743, *p*<.001) as well as information overload (*b*=.212, *p*<.010). Information overload, in turn, significantly drives frustration (*b*=.262, *p*<.010), whereas the direct path from BOL to frustration is insignificant. Thus, information overload fully mediates the method's direct impact on frustration (beyond its indirect effect via this path).

In contrast, there does not occur any mediation on user activity, neither via perceived novelty, nor via information overload – both paths are insignificant. There exists, however, a positive and significant direct effect from BOL to user activity (*b*=.390, *p*<.001).

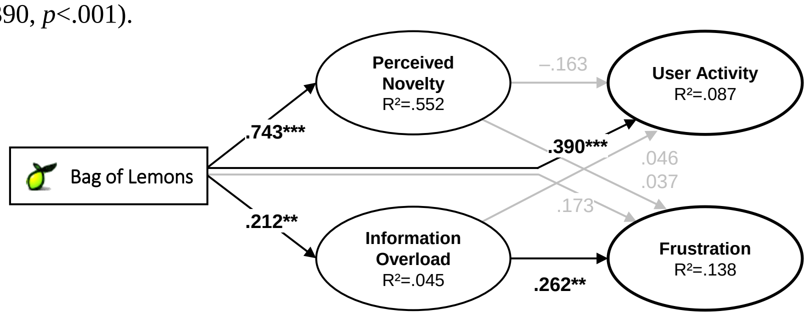


Figure 2. Structural Research model, including standardized path coefficients and R squared values (*** *p*<.001; ** *p*<.01)

Lastly, we considered the individual task completion times. Since this factor has an open-ended scale in one direction, Figure 3 depicts the main characteristics of the time distributions for the three treatment conditions in boxplot diagrams (indicating, median, as well as 25%- and 75%-quartiles). We find that the three conditions do not differ significantly in terms of completion time ($t_{BOL/Likert}$ = 1.564, *p*=.122; $t_{BOL/U\&D}$ = 1.467, *p*<.147; $t_{Likert/U\&D}$ = –.097, *p*=.923).

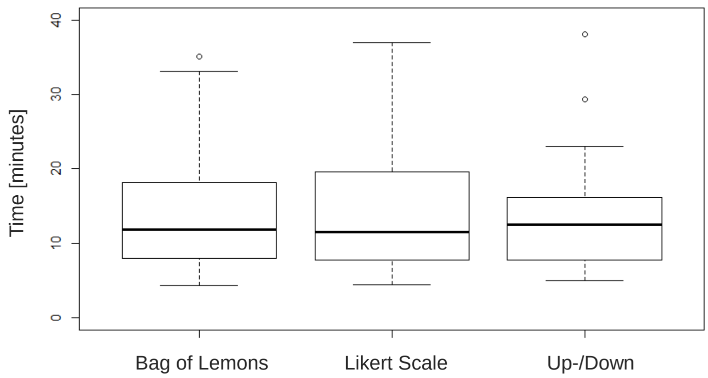


Figure 3. Boxplots of task completion times

5 Discussion and Conclusion

In this paper, we analyzed the effects of rating scales on users' activity, perceived information overload, perceived novelty, and frustration. In a field experiment in an open innovation campaign for a mid-size German research center, we assessed how BOL, up- and down-voting, and Likert scales differed in terms of these measures when employees were asked to evaluate a corpus of ideas created by their peers. All employees of the research center were invited to rate 42 proposals, being exposed to one of the above mentioned rating scales (between subjects design). Analyzing the behavioral as well as the post-evaluation survey data, we demonstrate that BOL, while stimulating activity, is also perceived as more frustrating than other rating techniques. We trace this result to the mediating factor of perceived information overload. Although participants were exposed to the same amount of information, that is, the identical corpus of 42 ideas, the bag of lemons method yielded much higher overload perceptions. We suggest that this may be due to deliberative and "pending" nature of the bag of lemons approach. While using Likert scales or up/down voting techniques, each idea can be assessed at a time, allocating lemons to a set of many ideas can be challenging since the desire to allocate a lemon late in the process may require to reassess previously rated ideas, for instance, to decide where to withdraw lemons from. This need for continuous cross-links requires to keep more ideas in mental "working memory," whereas they can be considered (and forgotten) sequentially when using the other techniques.

Coming back to our first research question of how the "Bag of Lemons" rating method affect user activity and frustration in a collaborative evaluation task, we hence can summarize that BOL increases both user activity and frustration. With regard to the second research question, that is, the role of perceived novelty and information overload in mediating these effects, we see that information overload fully mediates the effect of the BOL method on frustration, while perceived novelty does not exhibit any mediating properties. Moreover, there do not occur any cross-mediating effects, that is, from novelty to frustration or from information overload to activity.

Considering that approximately 50 percent of the employees of the institution evaluated their peers' proposals, this also hints at the high interest of employees in getting engaged in the process of participating in the decision-making processes at their workplace [28].

Theoretical and Practical Implications

This study contributes to the literature by evaluating a novel, thus hardly researched method of idea evaluation (BOL) in comparison to more established methods (Likert scales, up-/down voting). We focus on the important indicators of *user activity* and *frustration,* which has not or only scarcely been assessed by extant literature in this context. By integrating these opposing factors within a joint research model, we enhance the understanding of collaborative evaluation processes in view of differentiated rating regimes [5, 25, 26]. Next, by relating these key indicators to mediating factors, we provide starting points for understanding *how* the different rating methods affect the users' perceptions and behaviors. In particular, we identify perceived information overload as a potential mediating factor at play. Moreover, our study provides a show case of employee-driven innovation [27] and computer-supported organizational participation [28]. We confirm findings of Riedl et al. [20], who suggested that people form attitudes towards rating scales. Our findings also lend support to Klein and Garcia [7], underpinning BOL's novelty but, in contrast, do not confirm the method's superiority in terms of task completion time. Yet, we extend the authors findings by shedding light on users' perception of BOL's restraining character. Participants in our study expressed higher levels of frustration when evaluating ideas using the BOL as compared to Likert and up- and down-voting. This suggests that people might refrain from engaging in a BOL evaluation task in the future. Accordingly, practitioners should be aware of the possibly detrimental effects of BOL when designing an open innovation platform. This effect, as it is mediated by perceived information overload, may substantially be driven by the relatively high number of idea. We suggest that idea evaluation tasks with fewer ideas (e.g., 6 to 12), may yield different results.

Limitations and Future Research

Our study needs to be considered against several limitations. First, we compared the different rating methods in terms of user activity, frustration, and time, however, could not consider the evaluations' accuracy, that is, a match between the crowd's assessment versus how good the ideas actually were. This limitation points at several paths for future research, very much in the sense of prior studies [7, 20]. Future work needs to take into account accuracy, for instance by comparing the collaborative results with an expert rater panel.

Next, as we have shown in this study, BOL facilitates higher levels of (relative) user activity than other rating methods. Nonetheless, on average, Likert and up-/down votes yield a higher overall numbers of idea evaluations. Systematically varying the amounts of ideas and "lemons" to distribute could thus shed more light on the strengths and weaknesses of the BOL approach and its robustness against different set sizes.

Due to strict German data protection legislation at the workplace, we were only able to capture some demographic characteristics of our participants. Thus, the data set is somewhat incomplete and restricts us from fully taking into account potential age or gender effects. Based on the data we have, these characteristics did not differ between treatments, so that at least a treatment bias due to demographic factors could be ruled out. Another limitation relates to the fact that part of the correlation between the item-based measures may be due to common method bias as most data was collected using standard questionnaire items. User activity represents an exception; correlations here will not exhibit common method bias.

As this study finds rating scales to affect user frustration, we suggest that it is worth exploring the antecedents of scale-related techno-stress. The noteworthy differences for information overload between BOL and up- and down-voting already lend some support to this presumption.

Furthermore, our study as well as previous ones [7, 10, 15, 20, 34] asked participants to rate ideas in the absence of any indication on whether and how other users already rated proposals. Future research could thus investigate the impact of information cascades, that is, users being able to see the evaluations of other (earlier) users [56], which may significantly impact results [57]. Finally, future research should address how open innovation contests within companies shape employee commitment and overall productivity [28].

6 Acknowledgement

This study was part of the joint research project “Participation as a Service” (PaaS), funded by the German Federal Ministry of Education and Research.

References

1. Bjelland, O.M., Wood, R.C.: An Inside View of IBM’s “Innovation Jam.” MIT Sloan Manag. Rev. 50, 32–40 (2008).
2. Di Gangi, P.M., Wasko, M.: Steal my idea! Organizational adoption of user innovations from a user innovation community: A case study of Dell IdeaStorm. Decis. Support Syst. 48, 303–312 (2009).
3. Niemeyer, C., Wagenknecht, T., Teubner, T., Weinhardt, C.: Participatory Crowdfunding: An approach towards engaging employees and citizens in institutional budgeting decisions. In: 49th Hawaii International Conference on System Sciences (HICSS). pp. 2800–2808 (2016).
4. Chesbrough, H.W.: Open Innovation: The New Imperative for Creating and Profiting from Technology. (2003).
5. Adamczyk, S., Bullinger, A.C., Möslein, K.M.: Innovation Contests: A Review, Classification and Outlook. Creat. Innov. Manag. 21, 335–360 (2012).
6. Hossain, M., Islam, K.M.Z.: Ideation through Online Open Innovation Platform: Dell IdeaStorm. J. Knowl. Econ. 6, 611–624 (2015).
7. Klein, M., Garcia, A.C.B.: High-speed idea filtering with the bag of lemons. Decis. Support Syst. 78, 39–50 (2015).
8. Schultz, U., Vandenbosch, B.: Information Overload in a Groupware

Environment : Now You See It, Now You Don't. J. Organ. Comput. Electron. Commer. 8, 127–148 (1998).
9. Oldroyd, J., Morris, S.: Catching Falling Stars: A Human Resource Response to Social Capital's Detrimental Effect of Information Overload on Valuable and Visible Employees. Acad. Manag. Rev. 37, 396–418 (2012).
10. Blohm, I., Riedl, C., Füller, J., Leimeister, J.M.: Rate or Trade? Identifying Winning Ideas in Open Idea Sourcing. Inf. Syst. Res. 27, 27–48 (2016).
11. Teschner, F., Rothschild, D.: Simplifying market access: A new confidence-based interface. J. Predict. Mark. 6, 27–41 (2013).
12. Martinez-Torres, M.R.: Content analysis of open innovation communities using latent semantic indexing. Technol. Anal. Strateg. Manag. 27, 859–875 (2015).
13. Teubner, T., Saade, N., Hawlitschek, F., Weinhardt, C.: It's only pixels, badges, and stars: On the economic value of reputation on Airbnb. In: Australasian Conference on Information Systems 2016 (2016).
14. Ebner, W., Leimeister, J.M., Krcmar, H.: Community engineering for innovations: the ideas competition as a method to nurture a virtual community for innovations. R&d Manag. 39, 342–356 (2009).
15. Riedl, C., Blohm, I., Leimeister, J.M., Krcmar, H.: Rating Scales for Collective Intelligence in Innovation Communities: Why Quick and Easy Decision Making Does Not Get it Right. In: Proceedings of the 31st International Conference on Information Systems, St. Louis (2010).
16. Negahban, S., Sewoong, O., Shah, D.: Iterative Ranking from Pair-wise Comparisons. 1–9.
17. YouTube: Five Stars Dominate Ratings, https://youtube.googleblog.com/2009/09/five-stars-dominate-ratings.html.
18. Marion K. Poetz and Martin Schreier: The value of crowdsourcing: Can users really compete with professionals in generating new product ideas? J. Prod. Innov. Manag. 1–31 (2012).
19. Kamis, A., Koufaris, M., Stern, T.: Management Information Systems Research Center, University of Minnesota. MIS Q. 32, 159–177 (2008).
20. Riedl, C., Blohm, I., Leimeister, J.M., Krcmar, H.: The Effect of Rating Scales on Decision Quality and User Attitudes in Online Innovation Communities. Int. J. Electron. Commer. 17, 7–36 (2013).
21. Venkatesh, V.: Creation of Favorable User Perceptions: Exploring the Role of Intrinsic Motivation. MIS Q. 23, 239–260 (1999).
22. Hwang, Y., Yi, M.Y.: Predicting the Use of Web-Based Information Systems: Intrinsic Motivation and Self-Efficacy. In: Proceedings of the 8th Americas Conference onf Information Systems. pp. 1076–1081 (2002).
23. Blohm, I., Riedl, C., Leimeister, J.M., Krcmar, H.: Idea Evaluation Mechanisms for Collective Intelligence in Open Innovation Communities: Do Traders Outperform Raters? In: ICIS 2011 Proceedings. pp. 1–24 (2011).
24. Likert, R.: A Technique for the Measurement of Attitudes. Arch. Psychol. 22, 1401–1455 (1932).
25. Leimeister, J.M.: Collective Intelligence. Bus. Inf. Syst. Eng. 4, 245–248

(2010).
26. Straub, T., Gimpel, H., Teschner, F., Weinhardt, C.: How (not) to Incent Crowd Workers: Payment Schemes and Feedback in Crowdsourcing. Bus. Inf. Syst. Eng. 57, 167–179 (2015).
27. Gressgård, L.J., Amundsen, O., Merethe Aasen, T., Hansen, K.: Use of information and communication technology to support employee-driven innovation in organizations: A knowledge management perspective. J. Knowl. Manag. 18, 633–650 (2014).
28. Wagenknecht, T., Filpe, R., Weinhardt, C.: Towards a Research Framework of Computer-Supported Organizational Participation. In: Tambouris, E., Panagiotopoulos, P., Sæbø, Ø., Wimmer, M.A., Pardo, T.A., Charalabidis, Y., Soares, D.S., and Janowski, T. (eds.) Electronic Participation: 8th IFIP WG 8.5 International Conference, ePart 2016. pp. 17–28. Springer Publishing (2016).
29. Wagenknecht, T., Crommelinck, J., Teubner, T., Weinhardt, C.: Ideate. Collaborate. Repeat. A Research Agenda for Idea Generation, Collaboration and Evaluation in Open Innovation. In: 13th International Conference on Wirtschaftsinformatik (2017).
30. Dean, D.L., Hender, J.M., Rodgers, T.L., Santanen, E.L.: Identifying Quality, Novel, and Creative Ideas: Constructs and Scales for Idea Evaluation. J. Assoc. Inf. Syst. 7, 646–699 (2006).
31. Sia, C.-L., Tan, B.C.Y., Wei, K.-K.: Group Polarization and Computer-Mediated Communication: Effects of Communication Cues, Social Presence, and Anonymity. Inf. Syst. Res. 13, 70–90 (2002).
32. Sassenberg, K., Postmes, T.: Cognitive and social processes in small groups : Effects of anonymity of the self and anonymity of the group on social influence. Br. J. Soc. Psychol. 41, 463–480 (2002).
33. Blohm, I., Bretschneider, U., Leimeister, J.M., Krcmar, H.: Does Collaboration among Participants Le ad to Better Ideas in IT-based Idea Competitions? An Empirical Investigation. In: Proceedings of the Annual Hawaii International Conference on System Sciences. pp. 1–10 (2010).
34. Bao, J., Sakamoto, Y., Nickerson, J. V.: Evaluating Design Solutions Using Crowds. In: Proceedings of the Americas Conference on Information Systems. p. Paper 446. , Detroit, Michigan (2011).
35. Fuchs, C., Schreier, M.: Customer empowerment in new product development. J. Prod. Innov. Manag. 28, 17–32 (2011).
36. Hawlitschek, F., Teubner, T., Weinhardt, C.: Trust in the Sharing Economy. Die Unternehmung. 70, 26–44 (2016).
37. Leimeister et al.: Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition. J. Manag. Inf. Syst. (2009).
38. Lakhani, K.R., Jeppesen, L.B.: Getting unusual suspects to solve R&D puzzles. Harv. Bus. Rev. 85, 30–32 (2007).
39. Poetz, M.K., Schreier, M.: The value of crowdsourcing: can users really compete with professionals in generating new product ideas? J. Prod. Innov. Manag. 29, 245–256 (2012).
40. Robinson, A.G., Schroeder, D.M.: Ideas are free: How the idea revolution is

liberating people and transforming organizations. Berrett-Koehler Publishers (2004).
41. Schwarz, N.: Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation. Lawrence Erlbaum, Hillsdale, NJ (1996).
42. Solomon, M., Bamossy, G., Askegaard, G., Hogg, M.K.: Consumer Behaviour: A European Perspective. Pearson FT Prentice Hall, Harlow, UK (2006).
43. Csikszentmihalyi, M.: Beyond Boredom and Anxiety. Jossey-Bass, San Francisco (1977).
44. Koufaris, M.: Applying the technology acceptance model and flow theory to online consumer behavior. Inf. Syst. Res. 13, 205–223 (2002).
45. Deci, E.L., Ryan, R.M.: Intrinsic Motivation Inventory, http://www.psych.rochester.edu/SDT/ measures/intrins.html.
46. Beaudry, A., Pinsonneault, A.: Understanding user responses to information technology: a coping model of user adaptation. MIS Q. 29, 493–524 (2005).
47. Maier, C., Laumer, S., Eckhardt, A., Weitzel, T.: Online Social Networks As a Source and Symbol of Stress: an Empirical Analysis. In: Proceedings of the 33rd International Conference on Information Systems, Orlando 2012 (2012).
48. Koroleva, K., Krasnova, H., Günther, O.: “STOP SPAMMING ME!” - Exploring Information Overload on Facebook. In: Proceedings of the 16th Americas Conference on Information Systems (AMCIS). p. Paper 447 (2010).
49. Eckhardt, A., Maier, C., Buettner, R.: The Influence of Pressure to Perform and Experience on Changing Perceptions and User Performance: A Multi-Method Experimental Analysis. Proc. 33rd Int. Conf. Inf. Syst. 1–12 (2012).
50. Haines, R., Hough, J., Cao, L., Haines, D.: Anonymity in Computer-Mediated Communication: More Contrarian Ideas with Less Influence. Gr. Decis. Mak. Negot. 23, 765–786 (2014).
51. Wagenknecht, T., Teubner, T., Weinhardt, C.: The Impact of Anonymity on Communication Persuasiveness in Online Participation. In: Proceedings of the Thirty Seventh International Conference on Information Systems (ICIS) (2016).
52. Galletta, D.F., Henry, R.M., Mccoy, S., Polak, P.: Web site delays: How tolerant are users? Inf. Syst. Res. 17, 20–37 (2002).
53. Ringle, C.M., Wende, S., Becker, J.-M.: SmartPLS 3. , Bönningstedt (2015).
54. Chin, W.W.: The partial least squares approach to structural equation modeling. Mod. Methods Bus. Res. 295, 295–336 (1998).
55. Gefen, D., Straub, D.W., Boudreau, M.-C.: Structural equation modeling and regression: Guidelines for research practice. Commun. Assoc. Inf. Syst. 4, 1–77 (2000).
56. Bikhchandani, S., Hirshleifer, D., Welch, I.: A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. J. Polit. Econ. 100, 992–1026 (1992).
57. Duan, W., Gu, B., Whinston, A.B.: Informational Cascades and Software Adoption on the Internet: An Empirical Investigation. MIS Q. 33, 23–48 (2009).

The Rise of Crowd Aggregators
- How Individual Workers Restructure Their Own Crowd -

David Durward[1], Ivo Blohm[2]

[1] University of Kassel, Information Systems, Kassel, Germany
david.durward@uni-kassel.de
[2] University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland
ivo.blohm@unisg.ch

Abstract. Crowd work has emerged as a new form of digital gainful employment whose nature is still a black box. In this paper, we focus on the crowd workers – a perspective that has been largely neglected by research. We report results from crowd worker interviews on two different platforms. Our findings illustrate that crowd aggregators as new players restructure the nature of crowd work sustainably with different effects on the behavior as well as the existing relationships of crowd workers. We contribute to prior research by developing a theoretical framework based on value chain and work aggregation theories which are applicable in this new form of digital labor. For practice, our results provide initial insights that need to be taken into account as part of the ongoing discussion on fair and decent conditions in crowd work.

Keywords: Crowdsourcing, Crowd Work, Digital Work, Division of Work.

1 Introduction

In the last few decades the nature of work and employment relations has been changed sustainably on various levels, particularly caused by the restructuring of value chains [1]. As a result, the relationships are becoming unstable and the number of self-employed people is increasing in many industries [2]. With the rise of new information and communication technologies (ICT) as well as the internet, value chain restructuring relates to digitally mediated services. Nowadays, these services increasingly take place on online labor markets where labor is exchanged for money via the internet [3]. However, in course of the online labor markets a new form of digital work has emerged, i.e. *crowd work.*

This phenomenon can be described as a distinct type of labor that is located at the intersection of digital work and gainful employment, in which an undefined mass of people (i.e., crowd worker) creates digital goods via an open call [4]. Substantial parts of the value creation take place on IT-facilitated platforms provided by intermediaries. These intermediaries usually divide the tasks into discrete subtasks, distribute them and subsequently aggregate the contributions to a final solution [5].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Durward, D.; Blohm, I. (2017): The Rise of Crowd Aggregators - How Individual Workers Restructure Their Own Crowd -, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 395-409

Crowd work has shown a strong track record as the number of platforms and crowd workers has been growing continuously. Hence, the World Bank estimates the total crowd work market to be $4.8 billion in 2016 and up to $25 billion in 2020 [6].

Despite this rather growing importance, research on crowd work is still in its inception, in particular regarding the ones who perform the work, i.e. the crowd workers. Prior research that focused on the individual crowd worker have examined their motivations to participate in different types of projects [e.g., 7, 8], their demographical backgrounds [9], or analyzed antecedences of their task performance [10]. Furthermore, other researchers focused on characteristics of the crowd that meet specific organizational needs [11], as well as trust-related aspects [12]. Although few studies have been conducted to address the individual worker, there is a gap in understanding experiences and perceptions of crowd workers [13], in particular regarding the structure of work. Most research focused on the intermediaries and the processes of work aggregation on online labor markets [14, 15]. However, besides the ongoing value chain restructuring and its potentials, new hierarchies can be observed in online labor markets that apparently need to be analyzed out of an individual's view in order to gain a better understanding [1].

Thus, we examine aspects of restructuring and aggregation in crowd work context out of a crowd workers` perspective. For practice, it is essential to understand the continuous reshaping in crowd work in order to anticipate effects on the workforce. Furthermore, we contribute to the research fields of crowd work as well as online labor markets by describing more precisely the relationships between the involved parties and the prevailing conditions in the crowd-based value chain.

Therefore, we intend to fill the outlined research gaps regarding the perception of crowd work by addressing the following research question:

RQ: How do individual workers perceive the nature of crowd work regarding structure and aggregation?

2 Conceptual Background

2.1 Crowdsourcing and Crowd Work

The phenomenon of crowdsourcing describes a new form of outsourcing tasks, or more accurately, value creation activities and functions. According to Blohm et al. [16], the fundamental idea of crowdsourcing is that a crowdsourcer (which could be a company, an institution or a non-profit organization) proposes to an undefined group of contributors or crowd workers (individuals, formal or informal teams, other companies) the voluntary undertaking of a task presented in an open call. In this context, the ensuing interaction process unfolds over IT-based platforms. These platforms are provided by crowdsourcing intermediaries that assure the connection between the crowdsourcing companies (i.e. the crowdsourcers) and the crowd workers. Since these intermediaries provide platforms on which supply and demand of labor meet, they represent online labor markets as well as a new approach of work organization [17]. Figure [1] illustrates the traditional crowdsourcing context including the three mentioned parties.

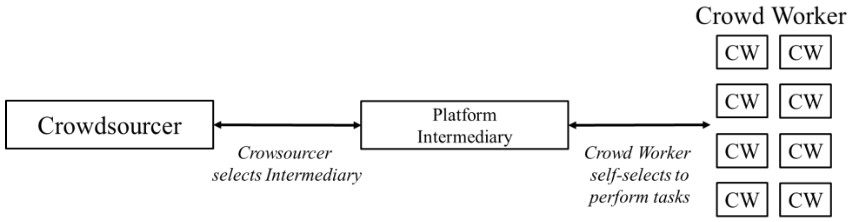


Figure 1: Roles and mediation in crowdsourcing initiatives (Source: adapted from Zogaj, Bretschneider et al. [18])

Furthermore, research has found important differences between the notions of crowdsourcing and crowd work [e.g., 4, 19]. According to Durward et al. [4] crowd work resembles a distinct type of labor that is located at the intersection of digital work and gainful employment. While crowd work is always paid, participation in crowdsourcing initiatives may have different motives and does not necessarily require financial remuneration, for example unpaid work that is done for a common good promoted by galleries, libraries, archives, or museums [e.g., 20]. Thus, out of an individual's perspective, crowd work reflects a kind of digital gainful employment that is based on crowdsourcing as organization principle. In this paper, we focus on the perception of crowd work out of an individual's view.

2.2 Value Chain Restructuring

In business studies literature, the value chain is an old-established concept that has been predominantly used by Porter [21] and describes a sequence of productive (i.e., value-added) activities leading to the delivery, consumption and maintenance of goods and services [22]. The term value chain is also used to emphasize the power relations, the vertical ties as well as sequential stages of production as well as service provision processes [23]. In this context, value chains are seen as dynamic and reconfigured on an ongoing basis [22]. However, online labor markets seem to be at the cutting edge of this new era of service chain value restructuring since they bring together buyers and sellers of digitally mediated service work [1].

Against this backdrop, crowd work is performed on online labor markets as crowd workers sell their skills and labor to crowdsourcers in order to generate various services that are mediated by the platform intermediary. In the last few decades, this service value chain restructuring has shaped the nature of work and employment sustainably [23]. In particular, the pattern of reintermediation has been shown to be a profound change in online labor markets since it refers to the disappearance of direct connections between clients and workers [1]. Against this backdrop, emerging structures, new forms of intermediation or aggregation in crowd work as a novel digital gainful employment need to be analyzed more precisely.

2.3 Work Aggregation

In general, aggregation is defined as the collecting of units or parts into a mass or whole [24]. This definition can apply to various contexts, including work and labor markets, on which work aggregators are able to break more complex projects into microtasks, distribute them to thousands of workers and subsequently aggregate the subtasks to a final solution [18]. A similar notion refers to online labor markets, in which work aggregators provide a managed service and platform usually as a layer on top of a intermediary's platform [25].

In crowd work context, there is an ideal-typical process of projects. Initially, the general task gets decomposed, described in detail and distributed to the crowd [4]. Breaking down tasks into subtasks can be provided either by the intermediary or the crowdsourcer itself. There are specific task modularization mechanism that provide functionalities that enable crowdsourcers to divide tasks into fine-grained subtasks [5]. Afterwards, the actual processing of the tasks takes place, before the solutions get selected and aggregated. Thus, this task decomposition and the reintegration are accomplished by the crowdsourcer in collaboration with the intermediary [5]. Against this backdrop, the analysis-synthesis concept comprises the decomposition of the main task into subtasks delegable to people and further the synthesis of subtasks results in order to reach the goal of the whole organization [26]. In previous research on crowd work [e.g., 5, 27, 28], work aggregation has mainly focused on the intermediaries or on the organizations as crowdsourcers.

However the perceptions and roles of the individual crowd workers regarding aggregation have been widely neglected. We assume that there are specific forms of aggregation within crowd work due to the predominant heterogeneity of potential contributors, the varying complexity of tasks and the asymmetric relationships between involved actors.

3 Research Method

3.1 Research Context and Data Collection

In order to develop our theoretical model, we analyzed the work context and perception of crowd workers on two crowdsourcing platforms – i.e., *Elance.com*[1] and *Freelancer.com.*[2] With the aim of preventing elite bias [29], we have chosen these marketplaces to overcome biases resulting from a single intermediary and due to the diversity of offered tasks as well as the various types of crowd workers. The primary data source contains of 12 semi-structured interviews since this kind of interviews are well suited in exploring attitudes, values, beliefs as well as the views of a person towards a phenomenon of interest [30]. Hence, we decided to conduct semi-structural interviews to understand the socio-technological context of crowd work out of a workers' perspective and extract their individual views regarding the nature of work.

[1] https://www.upwork.com/about/
[2] https://www.freelancer.com

In developing the interview protocol, we therefore used Kvale's [30] framework of conversational, qualitative interviewing as a template to ensure that our semi-structured interview elicit information relevant to our research question. Based on these guidelines, we designed an open-ended interview protocol that focused on the work environment and the perception of work by using the well-established constructs of the work design questionnaire (WDQ) [31]. In IS research it is essential to provide an explicit framework for guiding the participants throughout the interview to articulate and interpret their experiences [32]. Since the key topics of the interviews derived from the WDQ as our framework, we had to modify the wordings of the questions and adjust them to the study context of crowd work. The interviews took place between December 2015 and January 2016. Every single interview lasted between 60 and 90 minutes and was conducted with voice over IP (VoIP) communication via *Skype*. Since a respondent and an interviewer might be less engrossed in an interview conducted by telephone than in person [33] and anonymity is assured, we therefore aimed to prevent interviewer bias as well as social desirability and thus generating accurate information [34]. Subsequently, the interviews has been transcribed, coded and analyzed by using the analysis software package *ATLAS.ti*.

Since we aim to provide an unbiased data basis, numerous and knowledgeable respondents who view the focal phenomena from diverse perspectives, have been interviewed. Thus, we select respondents who differ regarding the duration of marketplace membership and performed jobs, (e.g., designing, or coding). As the evaluation of perception and behavior could differ concerning their previous experience, we interviewed more and less experienced crowd workers. Furthermore, we also analyzed their personal data that was available on their publicly visible user profiles. It has been found that the duration of membership (i.e., the time registered on the given platform), the amount of clients, the number of performed jobs and the average hourly rate of the crowd workers are reliable indicators of experience (see Table [1]). We interviewed six crowd workers per intermediary.

Table 1. Selection of crowd workers

Crowd worker	*Membership*	*Category*	*Jobs*	*Average Hourly Rate*	*Clients*
CW1	Nov 11	Writing	64	$ 15	37
CW2	Nov 11	Writing	21	$ 27	15
CW3	Aug 13	Translation	53	$ 30	35
CW4	Nov 14	Programming	11	$ 11	6
CW5	Dec 12	Administration	16	$ 14	12
CW6	Jan 14	Programming	7	$ 15	5
CW7	Nov 07	Translation	272	$ 30	196
CW8	Jan 15	Writing	17	$ 25	16
CW9	Feb 14	Translation	18	$ 20	12
CW10	Apr 11	Writing	31	$ 30	28
CW11	May 14	Programming	24	$ 18	21
CW12	Aug 03	Programming	16	$ 80	13

3.2 Data Analysis

We want to find out how crowd workers behave and how they organize their work on crowdsourcing platforms. Thus, according to several researchers [e.g., 35], we apply the approach of Gioia, et al. [36] to analyze our qualitative data. This methodology basically consists of two separate analysis phases. In a first iteration, the analysis follows interviewee-centric terms and concepts in an inductive fashion (1st-order analysis). Within the phase of the 1st-order analysis, a myriad of terms, codes and concepts emerged in the analysis process. Looking for similarities and relations among the many codes we reduced the number of codes to a manageable amount by relating them to concepts. We tied to focus on concepts and tentative relationships emerging from the interviews in order to develop a comprehensive compendium of 1st-order terms [36]. In this context, concepts are vaguely specified notions that capture basic qualities of a phenomenon [36]. In a second step, we organized the 1st-order concepts into 2nd-order (theory-centric) themes and distilled them into overarching theoretical dimensions. These emerging 2nd-order themes indicate concepts that might help to explain the observed phenomena. Subsequently, we distilled the 2nd-order themes even further into aggregate dimensions [36].

In sum, having the 1st-order concepts, the 2nd-order themes and the aggregate dimensions, the foundation for building a data structure is provided. Besides its visualization, the data structure represents a presentation of the process from raw data to terms and themes in conducting the analysis and thus is an essential part of demonstrating rigor in qualitative research [37]. We then formulated dynamic relationships among the 2nd-order concepts in the data structure and transformed these insights into a theoretical model [36]. The focus of building models is how to account for not only all the major emergent concepts, themes and dimensions, but also for their dynamic interrelationships [36]. Against this backdrop, we want to find out how crowd workers perceive and organize themselves in crowd work as online labor markets by following this introduced approach.

4 Results

In a first step, we therefore provide the essential groundwork for theory-building by developing the data structure. Our data structure includes 1st-order concepts that are significant to the crowd workers and 2nd-order themes that are extracted overarching themes. Finally, both iterations enabled us to assemble the aggregated dimensions.

4.1 Constitutive Elements of a Theoretical Model

Relocation of Value Creation. Our findings provide information about several aspects of activities abroad the intermediary's platform. First, certain crowd workers increasingly acquire more tasks on the platform than they can perform by their own. This is an intended action since they subsequently forward parts of the initial job to other crowd workers or even to an own platform external workforce. These persons are usually acquaintances, friends or even family members. We found evidence that certain tasks are given to siblings, partners or the own children by crowd workers: *"I'm looking for someone in order to handle the workload. I'm looking for an editor as good as I am. And my son just had started. At least I can rely on him." (CW7).*

Furthermore, these persons are to be entrusted with certain tasks regularly and thus represent a standing pool of external human resources. Identified motives for this relocation of value creation parts are trust-related aspects, reliability and more efficient interaction between the crowd workers and well-known external persons. Thus the composition of the crowd changes since external contributors get either hired by crowd workers or acquired as part of the existing crowd without necessarily register on the crowdsourcing-platform.

Second, we observed that existing boundaries between the platform-based crowd work and external activities become blurred. We observed that some crowd workers use the crowdsourcing platform only as an acquisition tool for attracting new crowdsourcers. Once the crowd workers have made the initial contact and completed first tasks, the follow-up business will be subsequently realized off-platform via different channels. In this context an interviewee stated: *"I would say that most of the business takes place offside the platform […] Most people finally use the platform in order to acquire clients." (CW3).*

In fact, the crowd workers proactive use various communication technologies like VoIP, E-mail, phone or virtual workplaces to interact bilaterally with the crowdsourcers instead of using the provided infrastructure of the platform intermediary. A major part of tasks and jobs are consciously processed beyond the platform sphere. In addition, all so-called after sales activities (e.g., customer service) for the crowdsourcers are independently managed by the crowd workers aside the platform. The aim is to develop long-term relationships to the crowdsourcers and simultaneously to save platform fees for both parties. Thus, the original idea of giving problems or tasks into the crowd and all steps of this service value creation will take place on platforms, must be critically questioned. We observed a contrary trend, in which particularly further business relationships develop beyond the platforms. Communication and information exchange as well as the actual task performance with the crowdsourcers take place in external settings of the crowd workers.

In sum, we identified two forms of value chain restructuring in crowd work. On the one hand, we observe an additional step in the crowd-based value creation since single crowd workers acquire tasks and further redistribute these tasks to external persons. One the other hand, the crowd workers are cutting out intermediation by interacting more directly with crowdsourcers and ignore the provided platform infrastructures as well as terms and conditions. This increase of disintermediation is originated from the crowd workers itself and reconfigures the existing service value chains. Hence, the value chain restructuring in crowd work is primarily based on the external relocation of certain value creation steps beyond the platform.

Emergence of New Hierarchy. In general, the decision-making process in crowd work is mutual. The crowdsourcers and the crowd workers are in certain negotiating situations, in which they exert bargaining power on each other through the platform. We observe this to be different when certain crowd workers undertake fundamental functions of the platform such as the management of tasks. Our analyses show that single crowd workers acquire larger tasks from the platform, decompose them into smaller subtasks and distribute these to other crowd workers. In this context, the single crowd worker takes over the governance and management of the subtasks. Furthermore, this mediating crowd worker predefines the conditions of the subtasks like payment, milestones and deadlines, based on the earlier agreements with the crowdsourcer. The other crowd workers, who perform the subtasks, do not have that much space to negotiate in this context since the general conditions have been set already. Thus, decisions are no longer being made bilaterally but by the mediating crowd worker. *"Then I realized that he himself was just a first intermediate step from another client." (CW08),* a hired crowd worker reported. In particular, unexperienced crowd workers who have not yet performed a lot of tasks, are implicitly dependent on these forwarded subtasks. They rely on information and specific input of the mediating crowd worker.

Furthermore, the conventional relationships between crowd workers and crowdsourcers depend on the scope and the type of tasks and thus vary considerably in crowd work. Nevertheless, we observe the relationships between the mentioned mediating crowd workers and the hired crowd workers, to be more long term oriented. The crowd workers who distribute the subtasks aim to develop long-term business relations to the task-performing crowd workers irrespective of the scope or type of task. One reason for this is a certain level of quality assurance since this closer business relationship permits a better control out of the mediating perspective. In the view of the performing crowd workers, this relation might be advantageous as well since they get tasks on a regular basis and thus a more secure income. *"They want to hire you more often and can save the fees for Elance." (CW7),* an interviewee said.

The division of labor is an essential characteristic of crowd work that varies regarding the different nature of tasks. Nevertheless, our analyses indicate that the mediating crowd worker acquires large projects and tasks on the platform, then decompose them and subsequently again broadcast the subtasks to the crowd. For example, the translation of a book chapter in English to French and German does not usually involve any division of labor. The crowd worker acting as an intermediary, however, acquires as well as decomposes the actual task of translation into several subtasks and then distributes them to other crowd workers he will hire. One crowd worker actually translates the text to French, while a second crowd worker will do the same in German. A third crowd worker subsequently proofread the translations before the mediating crowd worker will aggregate the subtasks into a final solution, submit it on the platform and thereby present it to the crowdsourcer. Since this trend is observable in various types of tasks, we can state that the division of labor tends to increase when those mediating crowd workers are present.

Therefore, with the rise of the introduced mediating actor, an additional element of the service value chain has emerged. Since dependencies between crowd workers shift and coordination as well as interaction becomes unilateral, we observed recent developments in crowd work towards more hierarchical structures.

Formation of Specialized Sub-Crowds. The single crowd workers who manage own tasks and projects apart the platform intermediary act as new intermediaries in crowd work contexts themselves. Thus they need their own standing workforce to expand and gain even more jobs. In order to achieve a competitive advantage, these mediating crowd workers, thus acquire their own specialized crowd. They proactively contacting other crowd workers based on their experience and skills to work for them. One interviewee stated: *"There are always many tasks you need certain specialists for and those are forwarded […] so when you are chef, you say that the cutting of onions is taken over by the assistant." (CW12).* The aim is to develop an own pool of expertise that is committed in the long-term by regularly forwarded tasks. Thus, these selections of workers represent sub-crowds that partially use the infrastructure of the platform but are managed by single mediating crowd workers.

Furthermore, these sub-crowds extend the own portfolio of the mediating crowd worker since new services, based on the crowd workers` skills, can be offered on the platform. Against this backdrop, the single mediating crowd worker wants his sub-crowd to be highly diverse concerning their capabilities. For example, a single crowd worker who offers the development of application programming interfaces (API) on the platform by itself begins to build up his own sub-crowd in the area of software development. Thus, the single crowd worker hires a specialist for agile software development methods and another backend developer who is proficient in different programming languages. Finally, a third crowd worker who is an engineer for the design of software products will be hired. The new intermediary manages all activities, subsequently consolidates and aggregates the single subtasks to a final solution. From a marketing point of view, the mediating crowd worker extended its own portfolio by forms of horizontal and vertical diversification.

Once the mediating crowd workers have built up their own workforce, some of them actually invest in their sub-crowds. They provide equipment (e.g., professional software to translate texts) and even share their own expertise and knowledge (e.g., by providing good own design templates). In line with this, one worker mentioned: *"So I bought him the transcription software F4 […] and he would not been able to afford the 50 Euros himself." (CW7).* Thus, the mediating workers act as some kind of mentor and develop their sub-crowds.

In sum, we outline our data structure in figure [2] which illustrates the 2nd-order themes on which we built our model of crowd work aggregators. These insights enabled us to develop a theoretical model of structures and concepts that emerged from the data. Hence, figure [2] represents first building blocks of a theory that have to be aligned and set in relation to each other in the next section.

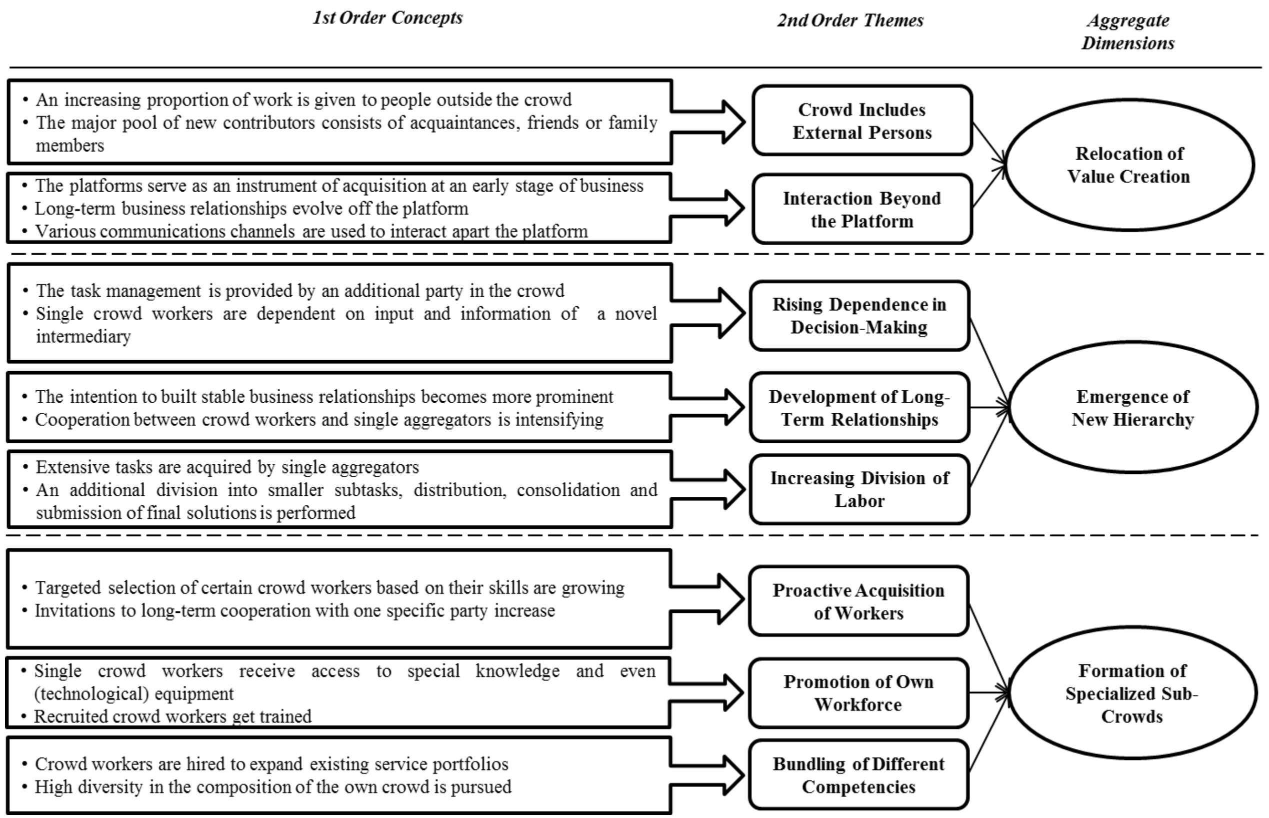


Figure 2: Data structure

4.2 The Role of Crowd Aggregators as New Elements of Value Creation

Although the data structure is essential, it is nonetheless the static picture of a dynamic phenomenon of interrelations [36]. Thus, we develop an inductive model that is grounded in the data of the crowd workers and captures the nature of crowd work in theoretical terms. Therefore, our model shows the dynamic relations amongst the emergent 2nd-order concepts which describe restructuring and reintermediation processes caused by new players in crowd work – i.e., the *crowd aggregators*.

The identified inclusion of external persons in existing service value creation processes changes the nature of crowd work sustainably. In particular, well-known persons of actual crowd workers are actively involved: *"Well, I have also invited friends and kind of activated them as a freelancer. They were previously no freelancers. And I knew what they are capable of and that they will work for that salary."(CW4).* In addition, the interactions between crowdsourcers and crowd workers increasingly take place not only via the platforms but using various channels aside. Hence, we propose that single crowd workers are responsible for this shift away from the actual platform-based crowd work towards a hybrid non-platform-based shape with external elements.

> **Proposition 1:** Crowd Aggregators consciously relocate parts of the value creation in crowd work.

Furthermore, certain crowd workers build up their own workforce by delegating prior decomposed subtasks to other crowd workers. In particular, unexperienced crowd workers are dependent on these tasks in order to gain reputation on the platform. Since the mediating worker usually unwinds repetitive tasks with the same crowd workers, this situation resembles an employer-employee-like relationship. Thus, a long-term relationship between the mediating crowd worker and the task-performing workers easily evolve. A mediating worker noted: *"In the end, I manage so to speak […] I develop myself towards project management and have my own sub-agency." (CW7*). Since the mediating crowd workers delegate and govern the work processes, we assume power asymmetries and dependencies to arise in these relationships. Hence,

> **Proposition 2:** Crowd Aggregators establish hierarchical structures in crowd work.

In addition, the aggregating crowd workers assemble a pool of other workers based on their capabilities and experience. These hired crowd workers further extent the own portfolio of the aggregator and thus represents a flexible workforce: *"I cannot program software. This is beyond my expertise. So, I hired another freelancer who is familiar with the technical details." (CW3).* Although the performing crowd workers are supervised, we found evidence indicating that the mediating crowd worker support its sub-crowd. On the one hand, the sub-crowd benefits from knowledge transfer with the aggregator and further its expertise. On the other hand, the mediating worker provides technical equipment if necessary and thus invest in the own specialized sub-crowd in order to gain reputation and generate more business itself on the crowdsourcing platform.

Proposition 3: Crowd Aggregators proactively develop own specialized sub-crowds.

This crowd aggregator represents a novel business idea in which the crowd is the core instrument of the service value chain. The aggregator is able to shape its own sub-crowd and act as a digital niche provider that guarantees efficient performance based on the specific composition of the sub-crowd. It exploits the provided infrastructure of the platform intermediary and reintermediates more specific services to the crowdsourcers. While the majority of the platform intermediaries focus solely on the pure mediation of tasks, the crowd aggregator supports the crowdsourcer during the entire procedure, particularly in larger and more complex projects. Out of the crowdsourcers view, the core competences of the crowd aggregator are coordination, decomposition and refinement of tasks as well as the final quality control of the single subtasks. Thus, the main assets of the crowd aggregator contain an efficient task assignment and successful management. In sum, we therefor denote a crowd aggregator as: *(1) An individual or a group of individuals that (2) act(s) as a novel intermediary and (3) use(s) existing platform infrastructure to build up its own specified sub-crowd.*

Figure [3] represents the core of our research results and outlines the data-to-theory connections. It illustrates the propositions (i.e., P 1-3) and their relations since it shows the two spheres (i.e., non- and platform-based), the composition as well as the involved parties of service value creation in crowd work. In addition, the model highlights the role and interrelations of crowd aggregators within crowd work context.

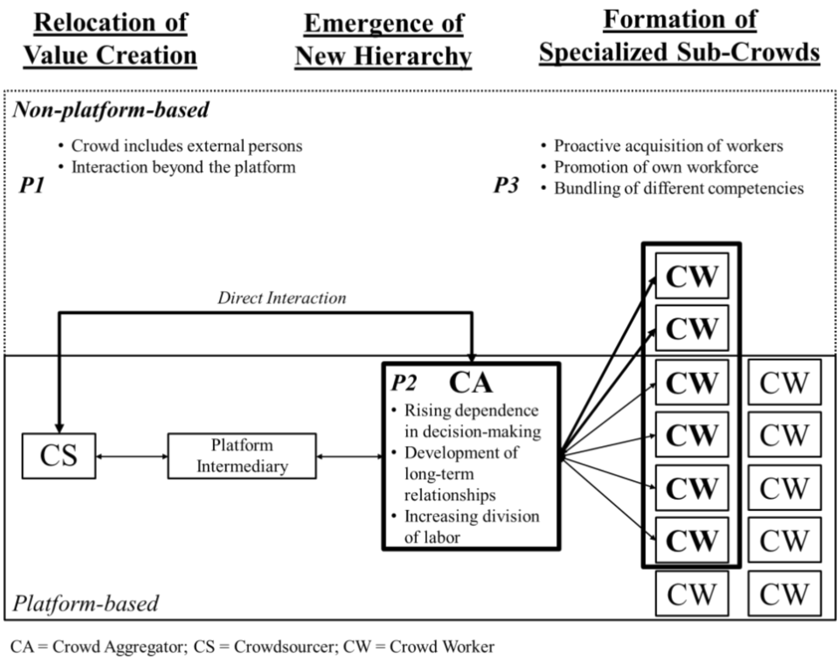


Figure 3: Model of crowd aggregators in crowd work

5 Discussion

We analyzed the nature of crowd work out of an individual's perspective. Thus, we address significant methodological shortcomings since we did not exclusively rely on online data about workers and their behavior [1]. In addition, our research contributes to literature on reintermediation that already takes place in online labor markets [e.g., 23], since we found evidence that the crowd aggregator represents more than an additional step in existing structures by illustrating their impact on structural as well as organizational level [1]. Previous research examined that crowd work describes a new system for the coordination of work that can be classified as ranking between the established forms of the two organizational principles of market and hierarchy [4]. Our findings indicate that crowd work redevelops towards hierarchical structures due to the rise of crowd aggregators. The unilateral decision-making processes, the long-term relationships between aggregator and its workers as well as the higher level of division of labor rather describe crowd work as more hierarchical. Furthermore, our findings can be explained by a lack of automated coordination mechanisms [38] in this crowd work context, since the crowd workers orchestrate the tasks themselves. Due to the missing of a higher-order coordination that provides the matching between the crowd and the offered tasks [26], crowd aggregators emerge as new coordinating elements. Our results are in line with the analysis-synthesis concept [26] since we illustrate that there is a need of certain decomposition and aggregation in crowd work.

For practice, platform intermediaries should closely monitor this development since the relocation of value creation aside the platform by the crowd aggregators may cause losses in control and income. In addition, our findings provide essential insights for the ongoing discussion about fair work conditions in the crowd. On the one hand, crowd aggregators exercise certain power over the crowd workers who are dependent on the aggregator. On the other hand, work conditions might be less precarious due to the rather long-term relationships, the constant supply of tasks and the enhanced provision of information. Nevertheless, the study has several limitations which constrain the generalizability of our results, since we developed our model gathering data from only two intermediaries. Further studies may overcome these limitations by evaluating the provided dimensions in subsequent empirical studies.

6 Conclusion

Given the lack of research on the individual in crowd work, our primary objective was to achieve a better understanding of the nature of work in the crowd. We followed a well-established methodology to conduct a qualitatively rigorous inductive study and developed a theoretical model of crowd aggregators. Our results illustrate that these crowd aggregators represent new players that restructure the workflows in crowd work on different levels and have impact on the behavior and relationships of the crowd workers. The crowd aggregators relocate activities off the platforms, reintermediate existing processes and build up an own sub-crowd. As a result, with their rise, crowd work evolves into a more hierarchical form of labor.

References

1. Lehdonvirta, V., Hjorth, I., Graham, M., Barnard, H.: Online Labour Markets and the Persistence of Personal Networks: Evidence From Workers in Southeast Asia. ASA Annual Meeting, Chicago, IL, 22nd Aug, (2015)
2. Hollister, M.: Employment stability in the US labor market: Rhetoric versus reality. Annual Review of Sociology 37, 305-324 (2011)
3. Horton, J.J.: Online labor markets. In: Conference Online labor markets, pp. 515-522. Springer, (Year)
4. Durward, D., Blohm, I., Leimeister, J.M.: Crowd Work. Business & Information Systems Engineering 58, 281-286 (2016)
5. Zogaj, S., Bretschneider, U.: Analyzing Governance Mechanisms for Crowdsourcing Information Systems - A Multiple Case Analysis. European Conference on Information Systems (ECIS), Tel Aviv, Israel (2014)
6. Kuek, S.C., Paradi-Guilford, C., Fayomi, T., Imaizumi, S., Ipeirotis, P., Pina, P., Singh, M.: The Global Opportunity in Online Outsourcing. (2015)
7. Brabham, D.C.: Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. Information, Communication & Society 13, 1122-1145 (2010)
8. Alam, S.L., Campbell, J.: Crowdsourcing motivations in a not-for-profit GLAM context: the Australian newspapers digitisation program. ACIS 2012: Location, location, location: Proceedings of the 23rd Australasian Conference on Information Systems 2012, pp. 1-11. ACIS (2012)
9. Kaufmann, N., Schulze, T., Veit, D.: More than fun and money. worker motivation in crowdsourcing–a study on mechanical turk. (2011)
10. Rogstadius, J., Kostakos, V., Kittur, A., Smus, B., Laredo, J., Vukovic, M.: An Assessment of Intrinsic and Extrinsic Motivation on Task Performance in Crowdsourcing Markets. ICWSM 11, 17-21 (2011)
11. Erickson, L., Petrick, I., Trauth, E.: Hanging with the right crowd: Matching crowdsourcing need to crowd characteristics. (2012)
12. Guo, W., Straub, D., Zhang, P.: The Impact of Formal Controls and Relational Governance on Trust in Crowdsourcing Marketplace: An Empirical Study. International Conference on Information Systems (ICIS), Milan, Italy (2013)
13. Deng, X.N., Joshi, K.: Is Crowdsourcing a Source of Worker Empowerment or Exploitation? Understanding Crowd Workers' Perceptions of Crowdsourcing Career. International Conference on Information Systems (ICIS), Milan, Italy (2013)
14. Quinn, A.J., Bederson, B.B.: Human computation: a survey and taxonomy of a growing field. CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 1403-1412 (2011)
15. Bederson, B.B., Quinn, A.J.: Web Workers, Unite! Adressing Challenges of online labores. Proceedings of CHI EA '11 CHI '11 Extended Abstracts on Human Factors in Computing Systems 97-106 (2010)
16. Blohm, I., Leimeister, J.M., Krcmar, H.: Crowdsourcing: How to Benefit from (Too) Many Great Ideas. MIS Quarterly Executive 4, 199-211 (2013)
17. Durward, D., Blohm, I., Leimeister, J.M.: Rags to Riches-How signaling behaviour causes a power shift in crowdsourcing markets. (2016)
18. Zogaj, S., Bretschneider, U., Leimeister, J.M.: Managing crowdsourced software testing: a case study based insight on the challenges of a crowdsourcing intermediary. Journal of Business Economics 84, 375-405 (2014)

19. Kittur, A., Nickerson, J.V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M., Horton, J.: The future of crowd work. Proceedings of the 2013 conference on Computer supported cooperative work, pp. 1301-1318. ACM (2013)
20. Alam, S., Campbell, J.: A conceptual framework of influences on a non-profit GLAM crowdsourcing initiative: a socio-technical perspective. 24th Australasian Conference on Information Systems (ACIS), pp. 1-11. RMIT University (2013)
21. Porter, M.E., Millar, V.E.: How information gives you competitive advantage. Harvard Business Review, Reprint Service (1985)
22. Sturgeon, T.J.: How do we define value chains and production networks? IDS bulletin 32, 9-18 (2001)
23. Flecker, J., Meil, P.: Organisational restructuring and emerging service value chains: implications for work and employment. Work, Employment & Society 24, 680-698 (2010)
24. Hackmann, H.: Governance theories and the practice of science policymaking. Science Policy 18 (2001)
25. Lehdonvirta, V., Ernkvist, M.: Converting the Virtual Economy into Development Potential: Knowledge Map of the Virtual Economy. infoDev / World Bank., Washington, DC (2011)
26. Buettner, R.: A Systematic Literature Review of Crowdsourcing Research from a Human Resource Management Perspective. HICSS, pp. 4609-4618 (2015)
27. Schenk, E., Guittard, C.: Crowdsourcing: What can be Outsourced to the Crowd, and Why? (2009)
28. Geiger, D., Seedorf, S., Schulze, T., Nickerson, R.C., Schader, M.: Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes. AMCIS, (2011)
29. Myers, M.D., Newman, M.: The qualitative interview in IS research: Examining the craft. Information and Organization 17, 2-26 (2007)
30. Kvale, S.: Interviews : an introduction to qualitative research interviewing / Steinar Kvale. Thousand Oaks, Calif. : Sage Publications. (1996)
31. Morgeson, F.P., Humphrey, S.E.: The Work Design Questionnaire (WDQ): developing and validating a comprehensive measure for assessing job design and the nature of work. Journal of applied Psychology 91, 1321 (2006)
32. Schultze, U., Avital, M.: Designing interviews to generate rich data for information systems research. Information and Organization 21, 1-16 (2011)
33. Rogers, T.F.: Interviews by Telephone and in Person Quality of Responses and Field Performance. Public Opinion Quarterly 40, 51-65 (1976)
34. Nederhof, A.J.: Methods of coping with social desirability bias: A review. European journal of social psychology 15, 263-280 (1985)
35. Mantere, S., Schildt, H.A., Sillince, J.A.: Reversal of strategic change. Academy of Management journal 55, 172-196 (2012)
36. Gioia, D.A., Corley, K.G., Hamilton, A.L.: Seeking qualitative rigor in inductive research notes on the Gioia methodology. Organizational Research Methods 16, 15-31 (2013)
37. Tracy, S.J.: Qualitative quality: Eight “big-tent” criteria for excellent qualitative research. Qualitative inquiry 16, 837-851 (2010)
38. Malone, T.W.: Modeling coordination in organizations and markets. Management science 33, 1317-1332 (1987)

Sind Smart Glasses die Zukunft der Digitalisierung von Arbeitsprozessen? Explorative Fallstudien zukünftiger Einsatzszenarien in der Logistik

Christina Niemöller[1], Benedikt Zobel[1], Lisa Berkemeier[1], Dirk Metzger[1], Sebastian Werning[2], Thomas Adelmeyer[1], Ingmar Ickerott[2] und Oliver Thomas[1]

[1] Universität Osnabrück, Informationsmanagement und Wirtschaftsinformatik, Osnabrück, Deutschland
{christina.niemoeller,benedikt.zobel,lisa.berkemeier, dirk.metzger,tadelmeyer,oliver.thomas}
@uni-osnabrueck.de
[2] Hochschule Osnabrück, Betriebswirtschaftslehre / Logistikmanagement, Osnabrück, Deutschland
{s.werning,i.ickerott}
@hs-osnabrueck.de

Abstract. Die Einführung von Smart Glasses eröffnet neue Chancen für die Gestaltung zukünftiger Arbeitsprozesse. Bisher sind diese Technologien wenig erforscht und werden nur experimentell hinsichtlich einzelner Aspekte untersucht. Zur Priorisierung zukünftiger Forschungsthemen und Identifikation relevanter Problemstellungen für den Bereich der Wirtschaftsinformatik wurden daher explorative Fallstudien mit zwei Logistikdienstleistern durchgeführt. Zur Ermittlung relevanter Einsatzszenarien wurde eine Triangulation aus Experteninterviews, Beobachtungen und Fokusgruppen gewählt und durch eine systematische Literaturrecherche ergänzt. Die 36 resultierenden Anwendungsfälle wurden mithilfe einer Umfrage priorisiert und auf Basis ihrer qualitativen Aussagen bzgl. der Herausforderungen analysiert. Die Ergebnisse des Beitrags sind (1) Einsatzszenarien für Smart Glasses in der Logistik sowie (2) daraus abgeleitete Forschungsthemen für die Wirtschaftsinformatik. Somit leistet diese Studie einen Beitrag zur Forschung im Bereich des ganzheitlichen Designs von Dienstleistungssystemen und zukünftiger Aufgaben digitaler Arbeit.

Keywords: Smart Glasses, Digitalisierung von Logistikprozessen, Fallstudie, Dienstleistungssysteme.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Niemöller, C.; Zobel, B.; Berkemeier, L.; Metzger, D.; Werning, S.; Adelmeyer, T.; Ickerott, I.; Thomas, O. (2017): Sind Smart Glasses die Zukunft der Digitalisierung von Arbeitsprozessen? Explorative Fallstudien zukünftiger Einsatzszenarien in der Logistik, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 410-424

1 Einleitung

Die Einführung von Smart Glasses verspricht neue Möglichkeiten zur Digitalisierung von Arbeitsprozessen. Vor allem in Situationen, in denen Mitarbeiter informationsintensive Tätigkeiten ausführen und gleichzeitig freie Hände benötigen, können Smart Glasses kontextsensitive Informationen in das Blickfeld der Nutzer einblenden und sie durch Arbeitsschritte leiten [1]. Die Logistik bietet aufgrund dieser Eigenschaften und zusätzlich starker Mitarbeiterfluktuation eine typische Anwendungsdomäne für Smart Glasses-basierte Dienstleistungssysteme.

Bisher sind diese Technologien wenig erforscht und werden nur experimentell hinsichtlich einzelner Aspekte untersucht (u.a. [2-4]). Eine systematische Literaturstudie zeigte, dass konkrete Anwendungsfälle in der Logistik bisher vor allem in praxisorientierten Fachmagazinen diskutiert werden (u.a. [5]). Dadurch bestätigt sich zum einen die betriebliche Relevanz des Themas, andererseits zeigt es die Lücke in der Forschung. Um dieser Lücke zu begegnen, wurde eine explorative Fallstudie in Kombination mit einer systematischen Literaturstudie durchgeführt, um Anwendungsfälle für die Logistik zu sammeln, abzugrenzen, zu definieren und weiteren Forschungsbedarf für die Wirtschaftsinformatik aufzuzeigen. Dabei wurden zum einen ein global agierender Logistikdienstleister mit Fokus auf die Kontraktlogistik sowie ein mittelständischer Logistiker mit Fokus auf die Fashion-Logistik ausgewählt. Zur Ermittlung relevanter Anwendungsfälle wurde eine Triangulation aus Interviews, Beobachtungen und Fokusgruppen gewählt und durch eine systematische Literaturrecherche ergänzt. Es zeigte sich, dass in der Literatur bisher vor allem klassische Themen wie Kommissionierung und Serviceunterstützung betrachtet wurden [6-8]. Die Fallstudie ergab allerdings zusätzliche Anwendungsfälle wie u.a. Workload Management, Integration der Kunden zur Direktabwicklung bspw. im Schadensmanagement (Value Added Services) oder automatisierte Prozessdokumentationen, welche bislang kaum betrachtet wurden. Zur Kategorisierung der Anwendungsfälle hinsichtlich weiterer Untersuchungen und zur Analyse ihrer Relevanz wurden die 36 ermittelten Anwendungsfälle mithilfe einer Umfrage bewertet und durch qualitative Aussagen zu möglichen Herausforderungen bei der Umsetzung ergänzt.

Die Ergebnisse des Beitrags sind Einsatzszenarien für Smart Glasses in der Logistik und eine Diskussion und Priorisierung dieser. Somit leistet dieser Artikel einen Beitrag zur Forschung im Bereich des ganzheitlichen Designs von Dienstleistungssystemen und zukünftiger Aufgaben digitaler Arbeit. Es liefert die Basis für weitere Forschung sowohl für (a) die gestaltungsorientierte Wirtschaftsinformatik, indem einzelne Szenarien mit Smart Glasses priorisiert zur Umsetzung vorgeschlagen werden und welche Herausforderungen dabei zu beachten sind (bspw. Datenschutz, technische und organisatorische Integration), als auch (b) für die verhaltensorientierte Forschung, u.a. welche Bereiche priorisiert betroffen sein werden und welche neuen Arbeitsformen sich ergeben. Der Beitrag gliedert sich wie folgt: In Abschnitt 2 wird zunächst das methodische Vorgehen begründet dargestellt. In Abschnitt 3 werden die abgeleiteten Anwendungsfälle aufgezeigt. In Abschnitt 4 werden diese hinsichtlich ihres weiteren Forschungsbedarfs diskutiert. Abschließend werden in Abschnitt 5 die Ergebnisse zusammengefasst und ein Ausblick gegeben.

2 Methodisches Vorgehen

2.1 Forschungsmethode

Um zu untersuchen, in welchen Prozessen Smart Glasses nutzbringend eingesetzt werden können, wurde eine Fallstudie im Rahmen des von Yin definierten Case Study Research Vorgehens gewählt [9]. Zwei Fälle innerhalb des Logistikbereichs werden unter der Benutzung von Replikationslogik evaluiert. Jeder Fall stützt sich auf mehrere Beweisquellen, aus denen Daten im Zeitraum der letzten 11 Monate extrahiert wurde.

Für die Fallstudie und die zu Grunde liegende Identifikation relevanter Use Cases wurden die Prozesslandschaften von zwei Logistikunternehmen untersucht. Die Auswahl der Unternehmen erfolgte auf Basis der Zielsetzung, möglichst viele unterschiedliche Prozesse mit hoher Stückzahl zwecks späterer Validierung zu untersuchen. Das ausgewählte global agierende Kontraktlogistikunternehmen (Fall A) zeichnet sich durch ein breites Spektrum an Anwendungsfällen aus dem Bereich der Lager- und Transportkonzepte aus. Diese reichen von See- und Luftfracht bis zur Lieferung per Zug und LKW mit mehr als 20 Mio. Sendungen pro Jahr. Ergänzt wird dies durch einen mittelständischen Logistikdienstleister (Fall B), der als Anbieter für Fashionlogistik vor allem als Experte im Bereich Kommissionierung und Value Added Services fungiert.

2.2 Datenerhebung und –auswertung zur Bestimmung der Use Cases

Zur Erhebung der Daten wurde eine Triangulation [10] aus (1) einer strukturierten Literaturrecherche, (2) Shadowing von Logistikprozessen, (3) Experteninterviews und (4) Diskussionen in Fokusgruppen gewählt.

Literaturrecherche. Zunächst wurde eine strukturierte Literaturrecherche durchgeführt [11, 12]. Dazu wurden folgende Datenbanken der Wirtschaftsinformatik und Wirtschaftswissenschaften (Ebscohost, ScienceDirect, AISeL, ACM, IEEE, Jstor, SpringerLink) sowie alle Journale des VHB Teilrankings der Logistik [13] analysiert. Zudem wurde die Suche auf die Datenbank WISO und Google Scholar ausgeweitet, um zusätzlich aktuelle Entwicklungen der Wirtschaft über tagesaktuelle Printmedien und sogenannte Grauliteratur abzudecken. Zur Identifikation von Use Cases in der Literatur wurde dazu die Kombination der Begriffe Datenbrille und Logistik, sowie deren englische Äquivalente, für die Suche verwendet (u.a. *"Smart Glasses" AND Logistics* und *Datenbrille AND Logistik)*. Die identifizierte Literatur wurde in Hinblick auf ihre Relevanz überprüft sowie entsprechend systematisiert und ausgewählt [12]: Einbezogen wurden Publikationen, (a) in denen Use Cases bzw. mögliche Anwendungsszenarien der Smart Glasses im logistischen Umfeld beschrieben werden und (b) in denen beispielhafte Umsetzungen von Smart Glasses Lösungen in Projekten des logistischen Umfelds erläutert werden. Ausgeschlossen wurden Publikationen (c) ohne Fokus auf die Logistikbranche, (d) mit thematischem Schwerpunkt auf „Smart Glass“ (intelligentes Fensterglas) und (e) in denen die Suchwörter ausschließlich getrennt voneinander behandelt werden. Es wurden 53 relevante Publikationen identifiziert.

Shadowing. Das Shadowing [10] diente im Rahmen der Forschungsarbeit dazu, durch Beobachtung der Arbeitsabläufe und Tätigkeiten erste potentielle Use Cases für Smart Glasses zu identifizieren. Zu diesem Zweck wurden jeweils am Standort Osnabrück eines globalen Logistikdienstleisters zunächst in einem Umschlagslager über zwei Tage alle relevanten Tätigkeiten zur Abwicklung des Stückgutgeschäfts analysiert und daraufhin in einem Distributionslager für drei Tage die Kernprozesse Wareneingang, Einlagerung, Kommissionierung und Warenausgang beobachtet. Die Erhebung wurde in Form von Feldnotizen und Prozessmodellen dokumentiert. Auf Basis der dokumentierten Prozesse wurden von den Wissenschaftlern die einzelnen Aktivitäten in Bezug auf potentielle Einsatzfelder für Smart Glasses diskutiert. Ein Use Case wurde definiert, wenn eine Aktivität durch Smart Glasses (a) entweder eine bessere Informationsbereitstellung oder (b) eine einfachere Umsetzung komplexer oder überflüssiger Arbeitsabläufe ermöglicht werden kann.

Experteninterviews. Durch Experteninterviews wird eine thematische Sondierung des Forschungsgegenstands aus einer praktischen Perspektive ermöglicht [14]. Im Rahmen der Experteninterviews wurden ein Business System Consultant sowie ein Business System Analyst mit 11 und 6 Jahren Berufserfahrung in der Position aus Fall A befragt. Beide Teilnehmer besaßen Expertise bezüglich der Informationssysteme (IS) und der Geschäftsprozesse des Logistikdienstleisters. Erste Kenntnisse über die Technologie der Smart Glasses lag insofern vor, dass die Probanden bereits Modelle wie die Google Glass selbst getestet, allerdings noch keine eigenen Implementierungen durchgeführt hatten. Der Leitfaden zur Durchführung der Experteninterviews bestand aus einer Reihe Fragen, die einen Transfer des Wissens der Experten über die logistischen Prozesse und die jeweiligen fachlichen Anforderungen auf die Technologie der Smart Glasses anregen sollten. Die Fragen waren dabei: (1) *Wie könnten Lagermitarbeiter im Umschlagslager durch Smart Glasses unterstützt werden?*, (2) *Haben Sie bereits mobile Informationssysteme im Güterumschlag im Einsatz?*, (3) *Wo würden Sie weitere Einsatzszenarien für Smart Glasses in der Logistik sehen?* und (4) *Gibt es Bereiche, Funktionen oder Tätigkeiten für die Sie den Einsatz von Smart Glasses kategorisch ausschließen würden?*.

Fokusgruppen. Die Durchführung der Fokusgruppen orientiert sich methodisch an dem von Bürki [15] und Misoch [16] beschriebenen Ansatz, indem zunächst Stimuli vorausgingen (hier Besichtigungen einzelner Läger von Fall A und B) und danach eine Gruppendiskussion durchgeführt wurde, dessen Ergebnisse von einem Protokollanten dokumentiert wurden: Es wurden vier Gruppendiskussionen an wechselnden Standorten durchgeführt. Der Kreis der Teilnehmer war überwiegend konstant, divergierte jedoch geringfügig auf Grund terminlicher Diskrepanzen und wechselnder Konsultation von externen Experten der Unternehmen A und B (u.a. Lagerleiter, Schichtleiter). Die konstanten Teilnehmer der Fokusgruppe bestanden aus (a) fünf Experten des logistischen Umfelds aus Fall A und B, (b) drei Wissenschaftlern aus dem Bereich Logistikmanagement, (c) zwei Experten aus der Software Entwicklung eines Software- und Beratungshaus und verknüpfend (d) vier Wissenschaftlern aus dem Bereich Dienstleistungsmanagement und Smart Glasses-Technologien (Leiter der Fokusgruppen-Diskussionen). Den Diskussionen gingen jeweils Stimuli in Form von Führungen durch ein Umschlags- und ein Liegewarenlager der Fälle A und B voraus.

Dies sollte als Impuls bzw. Anhaltspunkt für die Ideenfindung dienen und eine allgemeine Diskussionsgrundlage für die Teilnehmer schaffen. In den ersten drei Fokusgruppentreffen wurden jeweils mögliche Use Cases frei diskutiert, gesammelt und von Protokollanten dokumentiert. In der vierten Fokusgruppendiskussion wurden zu Beginn des Treffens die Summe der Use Cases aus der Triangulation (Literaturrecherche, Experteninterviews, Shadowing und den drei vorausgehenden Fokusgruppen) vorgestellt und anschließend weitere Use Cases ergänzt und diskutiert. Die Fokusgruppe diente neben der Sammlung von Use Cases (Fokusgruppentreffen 1-4) abschließend zur Validierung der aus den einzelnen Methoden aggregierte Liste der Anwendungsfälle. Hierzu fanden zwei weitere Fokusgruppen-Workshops statt (Fokusgruppentreffen 5-6). Dabei wurden auf Basis der aggregierten Liste aus der Triangulation (40 Use Cases) zum einen der Detailierungsgrad, die Verständlichkeit, die interne Konsistenz und die Anwendbarkeit zur Erreichung einer bereinigten und gültigen Liste der Use Cases [17] diskutiert. Dazu wurde jeder einzelne Use Case in Bezug auf Beschreibung, Akteure, Prozessphase (Wareneigang, Kommissionierung, Value Added Services etc.) und beispielhafte Instanziierungen des Use Cases diskutiert. Hierbei wurden die 40 Use Cases zu 36 Use Cases konsolidiert.

Datenauswertung. Zur Erfassung und Verdichtung der Use Cases wurde zunächst einzeln für jeden der vier Datensätze aus den vier Schritten der Triangulation die induktive Kategorienbildung nach Mayring [18] angewandt. Ein Use Case wurde definiert, wenn dieser entweder (a) eine funktionelle Unterstützung, (b) eine Umsetzung fachlicher Anforderungen der Logistik oder (c) potentielle oder erprobte Anwendungsszenarien beschreibt. Das heißt sobald eine der drei Kriterien zutraf wurde ein Use Case aufgenommen [18]. Wurde im weiteren Analyseverlauf wieder eine dazu passende Stelle gefunden, so wurde sie diesem Use Case ebenfalls zugeordnet (Subsumption), alternativ wurde ein neuer Use Case induktiv auf Basis des spezifischen Materials formuliert [18]. Eine grundsätzliche Herausforderung bei der Formulierung und Aggregation der Use Cases war das Auftreten teilweise stark divergierender Abstraktionsgrade bei der Beschreibung der Einsatzszenarien in den unterschiedlichen Methoden. Deshalb wurde, ausgehend von den Ergebnissen der vier verschiedenen methodischen Ansätze, erneut eine Kategorienbildung nach Mayring [18] angewandt, um die Use Cases schrittweise zu subsumieren. Bei der Durchführung dieser Vorgehensweise wurden genau solche Use Cases zusammengefasst, die eine möglichst hohe Homogenität bezüglich der Kurzbeschreibung und Zielführung aufwiesen. Abschließend wurden für alle Use Cases jeweils durch Domänenexperten der Fälle A und B eine Beschreibung (inkl. Akteure, Prozess und Aktivitäten) und Beispiele formuliert. Die Use Cases wurden in den Fokusgruppen-Treffen 5 und 6 validiert und konsolidiert (siehe „Fokusgruppen“) und bilden die Basis für die weitere Bewertung.

2.3 Datenerhebung und –auswertung zur Priorisierung der Use Cases

Datenerhebung. Im Rahmen der Fallstudie wurde eine personalisierte Online-Umfrage [19, 20] zur Bewertung der Use Cases durchgeführt. Die Bewertung der Use Cases wird ermittelt über den Grad der Übereinstimmung des Befragten mit einer vordefinierten Aussage. Die verwendete 7-stufige Skala umfasst die Werte 1 (Stimme gar nicht zu) bis 7 (Stimme stark zu), sowie „Nicht sinnvoll beantwortbar".

Der Fragebogen ist wie folgt strukturiert: Einleitend werden die 36 Use Cases in einer Übersicht dargestellt und als Orientierungshilfe zum Download zur Verfügung gestellt. Die Umfrage beginnt mit der Erfassung demographischer Daten, u.a. zum Arbeitsplatz und der Erfahrung im Umgang mit Smart Glasses. Anschließend erfolgt die Einschätzung der Use Cases einzeln, nach einem einheitlichen Fragenmuster. Dazu erfolgt zunächst die Beschreibung des Anwendungsfalls, unterstützt durch ein konkretes Beispiel im Logistikkontext (diese wurden jeweils durch Domänenexperten aus Fall A und B beschrieben und durch das Forschungsteam konsolidiert). Der Teilnehmer bewertet den jeweiligen Use Case im Anschluss anhand folgender vier Aussagen: (1) *Ich bewerte den Nutzen dieses Anwendungsfalls als hoch.* (2) *Ich bewerte den Innovationsgehalt dieses Anwendungsfalls als hoch.* (3) *Angenommen ich hätte Zugang zu Smart Glasses, dann würde ich sie für diesen Anwendungsfall einsetzen.* und (4) *Ich schätze den funktionalen/organisatorischen Aufwand der Einführung in meiner Organisation wie folgt ein.*

Die Bewertung dieser Aussagen ist obligatorisch für den Fortlauf der Umfrage und kann in einem freiwilligen Kommentarfeld ergänzt werden.

Definition des Fragenkatalogs. Existenziell für die Priorisierung von Umsetzungsprojekten ist die Auswahl relevanter Faktoren als Entscheidungsgrundlage [21]. Nach Jiang und Klein [27] sind für Projekte mit Technologien, deren strategische Auswirkungen langfristig bemerkbar sind, vor allem interne Einflussfaktoren (organisationale, anwenderbezogene, finanzielle, technische, und risikospezifische) entscheidend für die Projektwahl. In einem Fokusgruppen-Meeting wurden die Dimensionen zur Bewertung der Einsatzszenarien diskutiert und abschließend daraus vier Bewertungskriterien abgeleitet: (1) Organisatorischer Aufwand, (2) Nutzen, (3) Anwendungsbereitschaft und (4) Innovationsgrad. Der *organisatorische Aufwand* beschreibt den Aufwand für die Restrukturierung der Arbeitsprozesse. Für die organisationalen Ziele ist vornehmlich der *Nutzen* der Use Cases entscheidend. Die Fortschrittlichkeit der Use Cases und das damit verbundene Potenzial zur Restrukturierung werden durch den *Innovationsgrad* bewertet. Die *Anwendungsbereitschaft* der Use Cases wurde als anwenderbezogene und risikospezifische Dimension identifiziert. Die Business Use Cases repräsentieren die Anforderungen der Anwender; für eine *technische Bewertung* im Einzelnen, sind weitere Differenzierungen und Variantenbildungen in Form von System/IT-Use Cases erforderlich. Daher wurde in der Umfrage zunächst auf die Bewertung des technischen Aufwands verzichtet[1].

Beschaffenheit der Stichprobe. Die Stichprobe der Umfrage besteht aus einer bewussten Auswahl von Stakeholdern eines Smart Glasses-basierten Systems. Die Zielpopulation wurde nach dem Schneeballsystem konstituiert [20]: ausgehend von den Fokusgruppen wurde die Erhebungsgesamtheit auf betroffene Personenkreise innerhalb der beteiligten Institutionen erweitert. Es ergibt sich eine Grundgesamtheit von 31

[1] Eine Bewertung hinsichtlich des technischen Aufwands erfolgte im Nachgang durch Bildung einer Fokusgruppe aus acht IT-Experten und ist nicht Gegenstand dieses Beitrags.

Teilnehmern: 67 % Domänenexperten (10 Teilnehmer aus Fall A und 11 Teilnehmer aus Fall B), 7 % Implementierer und 26 % Wissenschaftler.

Datenauswertung. Basierend auf den mithilfe des Fragebogens erhobenen Daten wurden Datenwolken von den Kombinationen aus jeweils zwei Bewertungskriterien erstellt (Paarvergleichsmethode). Dieses Verfahren verhindert eine nachträgliche subjektive Beeinflussung der Ergebnisse (wie z.B. bei der Gewichtung oder Gruppierung von Kriterien) und erlaubt eine Fragmentierung der Gesamtfragestellung in weniger komplexe Einzelbetrachtungen [22]. Nach vorheriger Elimination nicht gültiger Werte (bspw. Auswahl „Nicht sinnvoll beantwortbar") wurde das arithmetische Mittel für jede der vier Kriterien (Nutzen, Aufwand, Anwendungsbereitschaft, Innovation) über alle Teilnehmer der Umfrage errechnet.

Für die Einteilung der Beobachtungen in einzelne Gruppen, bzw. Cluster, wurde das k-means-Verfahren als eines der populärsten Methoden des Clustering verwendet [23]. Der zugrundeliegende Algorithmus bestimmt die Zugehörigkeit der jeweiligen arithmetischen Mittelwerte zu einem Beobachtungscluster anhand der Entfernung der jeweiligen Beobachtungen zu einem der k-Clusterzentren. Dabei werden die k-Clusterzentren so verschoben, dass die kumulierten quadrierten Entfernungen aller Beobachtungen zum nächstgelegenen Clusterzentrum minimiert werden. Für die Bestimmung der optimalen Anzahl an Clusterzentren, k, wurde das Caliński-Kriterium verwendet [24].

Die qualitativen Daten wurden in ein Kategorienschema überführt, in dem gekennzeichnet wurde auf welches Statement oder welchen organisatorischen Aspekt sich der Kommentar bezieht. Die Aussagen wurden nummeriert mit A1-A9 und als Grundlage für die Ableitung von Implikationen herangezogen.

3 Use Cases für den Einsatz von Smart Glasses in der Logistik

Nach Anwendung der Methodik zur Verdichtung und Kategorisierung der Use Cases ergeben sich 36 Use Cases, welche nach ihrer jeweiligen Prozessschrittzugehörigkeit (Prozessgruppe) als auch ihrer Funktion (Funktionsgruppen) in der Fokusgruppe induktiv kategorisiert wurden. Als Prozessgruppen ergaben sich: *Management*, *Kommunikation*, *VAS*/*QS* und *phasenspezifische Use Cases* (*Wareneingang*, *Einlagerung*, *Kommissionierung*, *Warenausgang* und *Inventur)*. Ergänzend wurden folgende Funktionsgruppen definiert: *Monitoring*, *Steuerung*, *Anleitung*, *Video*, *Automatisierte Kontrolle*, *Identifizierung*, *Navigation* und *Sicherheit*. Abbildung 1 visualisiert die entstandene Landschaft aus 36 Anwendungsfällen[2] kategorisiert nach ihrer Prozess- und Funktionsgruppe.

Die Zuordnung eines Use Cases zu jeweils einer Prozessgruppe ist nicht in jedem Fall möglich, da gewisse Anwendungsfälle in mehreren Unternehmensprozessen vorkommen (n:m-Beziehung). Die Use Cases wurden daher der am nächsten stehenden Kategorie zugeordnet und mit der Fokusgruppe validiert. Für einige Gruppen ist ein Hilfs-Use Case identifiziert worden, welcher als Unterstützung für die Umsetzung der

[2] Eine detaillierte Beschreibung der 36 Use Cases kann unter https://www.imwi-data.uni-osnabrueck.de/Business-Use-Cases.pdf eingesehen werden.

anderen Business Use Cases der Gruppe dient. Diese Use Cases wurden in Abbildung 1 mit einem „H“ markiert. Es folgt eine Beschreibung der Funktionsgruppen und zur Veranschaulichung ausgewählten Use Cases (UC).

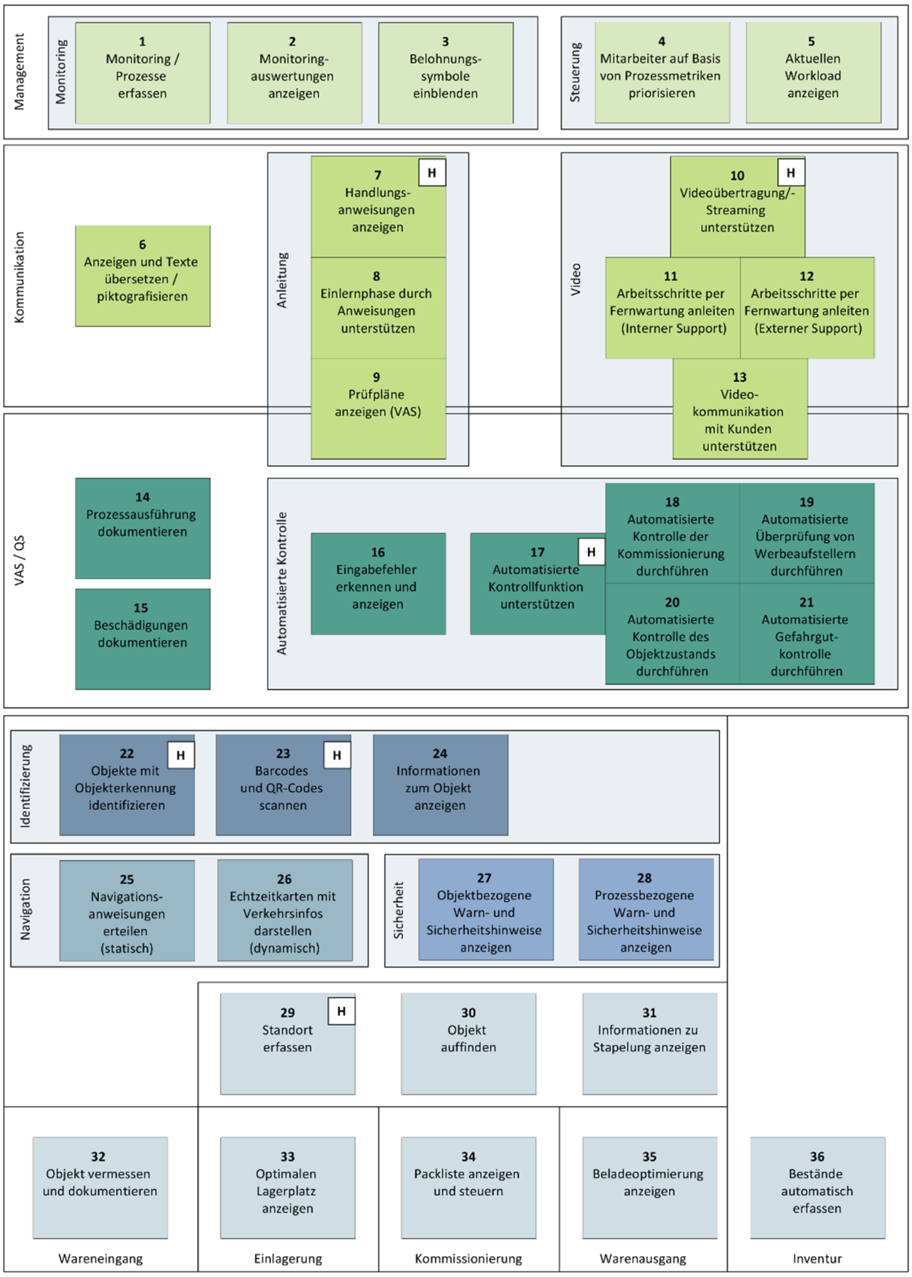


Abbildung 1. Kategorisierung der Use Cases

Management. Der Einsatz von Smart Glasses ermöglicht neue Use Cases zum *Monitoring* und *Steuern* von Mitarbeitern. Mithilfe der Smart Glasses können kontinuierlich Prozessdaten und Prozessabläufe während der Tätigkeit erfasst werden (UC1). Somit ist bspw. die Modellierung von Prozessen möglich, während der Mitarbeiter seine Tätigkeit durchführt. Mithilfe von Auswertungen (UC2) oder Belohnungssymbolen (UC3) kann dem Mitarbeiter z.B. das Erreichen einer Tagesleistung in der Kommissionierung und Prämien für weitere Leistung angezeigt werden. Auf Basis der angezeigten Auslastung (UC5) können mit Hilfe der Smart Glasses direkt einzelne Mitarbeiter gesteuert werden. Ein Schichtleiter kann bspw. kurzfristig einem Mitarbeiter auf dessen Smart Glasses eine priorisiert zugeordnete Aufgabe anzeigen (UC4).

Kommunikation. Die Funktionsgruppe *Anleitung* wird primär durch die Anforderung des Anzeigens von Handlungsanweisungen (UC7) als Hilfs-Use Case definiert. Dies kann z.B. in Form von Checklisten bei der Warenkommissionierung aber auch bei Prüfpläne für VAS-Dienstleistungen (UC9) zum Einsatz kommen. Für die Funktionalität *Video* ist im ersten Schritt eine Videoübertragung/Streaming notwendig. Ein bidirektionales Streaming ermöglicht Dritten, die Ansicht des Blickfelds des Smart Glasses-Trägers (z.B. für eine interne (UC11) oder externe (UC12) Fernwartung) zugänglich zu machen. Darüber hinaus ermöglichen die Smart Glasses eine vereinfachte Kommunikation durch die Übersetzung von Texten oder durch Piktografisierung (UC6), bspw. der Beschilderung für ausländische LKW-Fahrer.

Value Added Services (VAS) und Qualitätssicherung (QS). Der Einsatz von Smart Glasses ermöglicht eine *Automatisierte Kontrolle*. Dabei soll eine kamerabasierte Fehlererkennung und Rückmeldung an den Mitarbeiter (UC17) erfolgen, aber auch Fehler in der Eingabe erkannt und gemeldet werden (UC16). Besonders für die Qualitätssicherung sind eine Dokumentation, bspw. der nach den Service Level Agreements vereinbarten Prozessausführung, aus Sicht des Benutzers (UC14) als auch die Dokumentation von Beschädigungen mittels Kamera und Weiterleitung an das Schadensmanagement (UC15) potenzielle Einsatzszenarien.

Phasenspezifische Use Cases. Die Funktionsgruppe *Identifizierung* stellt eine Ansammlung von Schlüsselfunktionen dar. Hierzu gehören die Erkennung von Objekten anhand gespeicherter Merkmale (UC22), wie Farbe, Größe und Geometrie, als auch die Identifizierung mittels Bar- und QR-Code (UC23) sowie die Anzeige von Information zu den erkannten Objekten (UC24). Zur *Navigation* der Mitarbeiter bieten die Smart Glasses Möglichkeiten, zum einen statische Navigationsanweisungen, z.B. durch die Anzeige von Karten der Läger und Anweisungen zu erteilen (UC25), sowie eine dynamische Navigation (UC26) anhand von Echtzeitkarten mit Verkehrsinformationen anzuzeigen. Mithilfe der dynamischen Navigation könnten bspw. in einem Umschlagslager andere Staplerfahrer angezeigt werden, um Unfälle zu vermeiden. Einsatzszenarien im Bereich *Sicherheit* sind die Anzeige von objektbezogenen Warn- und Sicherheitshinweise (UC27), z.B. die Bruchgefahr von einzelnen Artikeln, wie auch prozessbezogene Sicherheitshinweise (UC28) bei z.B. dem Erreichen der maximal erlaubten Arbeitszeit. In beiden Fällen können akustische, optische oder haptische Signale für die Warnmeldung verwendet werden. Neben den funktionszugehörigen Use Cases stellt besonders der Anwendungsfall der systemunterstützen Standort-Erfassung (UC29) einen zentralen Hilfs-Use-Case dar. Weiterführend gibt es Use Cases, welche sich lediglich anhand von Standard-Lager-Prozessen, wie dem Wareneingang und der Kommissionierung kategorisieren lassen. Dazu gehört die Objektvermessung mittels Bilderkennung und spätere Dokumentation (UC32) aber auch die Packlisten Darstellung und Steuerung (UC34), welche die Verwaltung und die Abwicklung von Kommissionier-Aufträgen ermöglichen.

4 Implikationen für die Forschung

4.1 Analyse der Use Cases

Zur Priorisierung erster Use Cases zur Umsetzung und Ermittlung damit verbundener Forschungsbereiche für die Wirtschaftsinformatik wurden mithilfe der Paarvergleichsmethodik des Clustering (vgl. 2.3) erste Use Cases ausgewählt. Im Rahmen der Fokusgruppenworkshops haben die Vertreter aus Wissenschaft und Industrie den Kriterienvergleich von *Nutzen und Innovation* als Grundlage zur Identifikation besonders relevanter und forschungsorientierter Use Cases mit Innovationscharakter selektiert (siehe Abbildung 2).

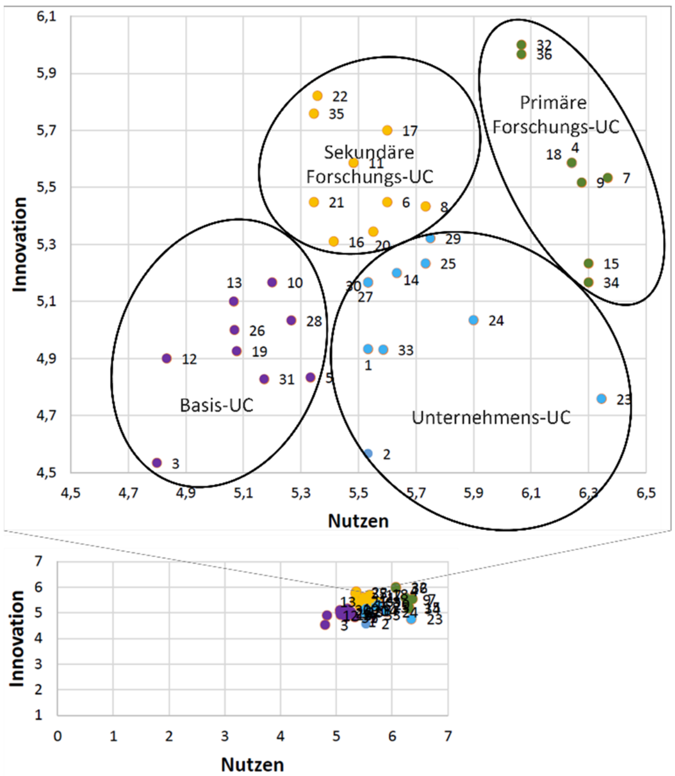


Abbildung 2. Kriterienvergleich Innovation zu Nutzen

Bei Betrachtung von Innovationsgehalt und Nutzen (vgl. Abbildung 2) ergeben sich vier Cluster: Use Cases mit vergleichsweise mittlerem Nutzen und mittlerem Innovationsgehalt (*Basis-UC*) sowie mit vergleichsweise hohen Nutzen aber nur mittlerem Innovationsgehalt (*Unternehmens-UC)* sowie mit einem vergleichsweise hohem Innovationsgehalt und mittlerem Nutzen (*Sekundäre-Forschungs-UC*). Die *Primären-Forschungs-UC* sind durch einen hoch bewerteten Nutzen und einen mittleren bis sehr hohen Innovationsgrad höchst relevant. Auf Basis dieser Auswertung ergeben sich acht Use Cases, die zunächst für weitere Forschungsaktivitäten priorisiert werden. Die quantitative Analyse wird durch eine Auswertung der qualitativen Daten aus den Kommentarfeldern der Umfrage vervollständigt. Eine Übersicht der priorisierten Use Cases (A0) und den zugehörigen qualitativen Aussagen wird in Tabelle 1 gegeben (Aussagen A1-A9). Die Reihenfolge der Use Cases folgt der Häufigkeit der Nennung in den relevanten Clustern.

Tabelle 1. Priorisierte forschungsrelevante Use Cases (A0) und qualitative Aussagen (A1-A9)

Use Case	*Anmerkungen*	*Nr.*
(4) Mitarbeiter auf Basis von Prozessmetriken priorisieren	Erhöhte Anforderung an vorgelagerte Systeme (ERP, LVS, PPS, MES)	A1
	Rechtliche Bedenken (Betriebsrat, Datenschutz)	A2
	Prozessveränderung nach Umsetzung	A3
(7) Handlungsanweisungen anzeigen	Erhöhte Anforderung an vorgelagerte Systeme (ERP, LVS, PPS, MES)	A1
	Neue Anforderung von zu pflegenden Daten	A4
(9) Prüfpläne anzeigen (VAS)	Analog Potential für weitere Value Added Service (VAS) Prozesse eruieren	A5
	Integration dynamischer Prozessveränderungen erforderlich	A3
(15) Beschädigungen dokumentieren	Erhöhter Aufwand bei technischer Umsetzung	A6
(32) Objekte vermessen und dokumentieren	Technische Umsetzung nicht trivial je nach Objektgröße	A6
(36) Bestände automatisch erfassen	Heute vorhandenes Potential der Bilderkennung ist zu prüfen	A6
	Technische Umsetzung ist aufwändig	A6
	Hohe Anforderung an die verbaute Technik in den Smart Glasses	A6
	Realer Nutzen abhängig vom Objekt und der Umgebung - ggfs. nur für ausgewählte Warengruppen möglich	A7
(18) Automatisierte Kontrolle der Kommissionierung	Organisatorischer Aufwand (Einmaliger Aufwand für Lagerfachkennzeichnung)	A8
(34) Packliste anzeigen und steuern	Realer Nutzen abhängig vom Objekt und der Umgebung	A7
	Erhöhte Anforderung an vorgelagerte Systeme (ERP, LVS, PPS, MES)	A1
	Smart Glasses als Visualisierungswerkzeug	A9

Als Validierung und Ergänzung wurden die im Rahmen der Umfrage erfassten Kommentare ausgewertet. In drei Fällen (UC 15, UC 32, UC 36) wurde die *technische Umsetzung* als nicht trivial angesehen sowie bei drei Anwendungsfällen (UC 4, UC 7, UC 34) eine erhöhte Komplexität bzgl. der *IT-Architektur* und Anbindungen an vorgelagerte Systeme (ERP, LVS, PPS, MES) gesehen. Weiteren Forschungsbedarf gibt es, den qualitativen Kommentaren der Umfrage folgend, vor allem in Bezug auf *Datenschutz* und *organisatorischer Integration* (u.a. Prozesse und Datenhaltung).

4.2 Diskussion

Die in der Fallstudie ermittelten Ergebnisse bzgl. einer Priorisierung der Umsetzung einzelner Use Cases sowie weiterer Forschungsthemen bieten Implikationen für die Wirtschaftsinformatik. (A0) Für die Forschung besonders interessante Use Cases (hoher Innovationsgrad bei gleichzeitig hohem Nutzen) können Tabelle 1 entnommen werden. Zur Umsetzung dieser Use Cases sind laut gestaltungsorientierter Forschung zunächst die Meta-Anforderungen für die konkreten Einsatzszenarien und Design-Prinzipien zu spezifizieren [25, 26]. Die systematische Literaturrecherche zeigte, dass bisher eine geringe Literaturbasis für Smart Glasses-basierte Dienstleistungssysteme besteht. Daher stellt sich im Besonderen die Frage *Auf welchen Bereichen der Knowledge Base kann zur Gestaltung der Design Prinzipien von Smart Glasses-Systemen für die Logistik aufgebaut werden?* Die Use Cases und deren qualitative Anmerkungen zeigen, dass vor allem (A1, A3) Prozessorientierung und (A9) Visualisierung als wichtige Themen erachtet werden, sodass Analogien aus diesen Bereichen bspw. aus dem Wearable Computing Design (u.a. [27-29]) und der Untersuchung von mobilen prozessorientierten Assistenzsystemen aus der Dienstleistungsforschung (u.a. [30-32]) herangezogen werden können. Analysen der technischen Funktionen von Smart Glasses [1] (A6) können darüber hinaus Implikationen für die Design Prinzipien sowie Integrationsfragestellungen geben.

Bezüglich der Integration in die IT-Architektur (A1, A6) ergeben sich Fragen aus den technischen Eigenschaften der Smart Glasses, bspw. in Bezug zu den Sensoren, und wie diese in die Schnittstellen der bestehenden Architektur integriert werden können. Die Unternehmensarchitektur und die Geschäftsprozesse betreffend stellen sich u.a. folgende Fragestellungen, die sich auch in den qualitativen Aussagen widerspiegelten: *Welche Veränderungen ergeben sich in den Geschäftsprozessen durch die Umsetzung der Use Cases?* (A3) Bzw. *Ergeben sich durch den Einsatz von Smart Glasses neue Arbeitsmuster und –arbeitsformen?* (A3) Bspw. durch den Use Case „Bestände automatisch erfassen“ (UC36) können interne Inventurprozesse neugestaltet bzw. komplett durch eine kontinuierliche Erfassung der Prozesse ersetzt werden. Gleiches gilt für den Use Case „Beschädigungen dokumentieren“ (UC15). Hier zeigte das Shadowing, dass bisher ein aufwendiger Retourenprozess ablief, mit nachträglichem Erfassen von Schäden und dem Zeitpunkt des Eintritts. Durch die Smart Glasses kann bei jedem Prozessschritt (bspw. Wareneingang, Umschlag, Auslagerung) kontinuierlich ein Foto des Objektes zur einfacheren Dokumentation erstellt werden. Hier ergeben sich weiterführende Anforderungen an das Datenmanagement (A4). Ein großer Forschungsbereich, der sich aus den diskutierten Einführungen ergibt, ist der Datenschutz (A2). Hinsichtlich der Berücksichtigung in der Systementwicklung ergibt sich zunächst die Forschungsfrage: *Welche datenschutzrechtlichen Anforderungen an Smart Glasses bestehen und wie kann ein datenschutzkonformes System gestaltet werden?* In der Literatur wird die isolierte Betrachtung von Datenschutzaspekten bei der Entwicklung von Privacy-invasiven Systemen (so auch Smart Glasses) kritisiert und die Integration einer Anwenderevaluation in die Systemkonstruktion gefordert [33]. Somit stellt sich in Bezug auf die konkrete Entwicklung eines Prototyps weitergehend die Forschungsfrage, wie Maßnahmen der Informationsprivatheit in die Systementwicklung integriert werden können. Zur Begegnung der geäußerten Bedenken hinsichtlich des Betriebsrats (A2) ist, einhergehend mit dem Datenschutz, die Akzeptanzfrage zu betrachten. Es ist zu untersuchen, ob in der Literatur bekannte Modelle den erweiterten Funktionsumfang neuer IKT erfassen [34] oder, ob die Integration neuer Einflussgrößen erforderlich ist (vgl. [35]). Die Akzeptanz betreffend können ggf. Analogien aus dem Wearable System Engineering herangezogen werden, bspw. zur Verbesserung von Arbeitsschutz und Ergonomie [36]. Abschließend wurde in der Studie die Frage nach Übertragungspotenzialen auf weitere Szenarien (A5, A7, A8) angemerkt. Dies impliziert für die weitere Forschung, dass einzelne Use Cases kontinuierlich sowohl summativ als auch formativ während der Gestaltung und Implementierung zu evaluieren sind, wie es auch aktuell in der gestaltungsorientierten Forschung diskutiert wird [37, 38].

5 Fazit

Bisher gibt es wenig wissenschaftliche Literatur bzgl. des Designs und der Implementierung Smart Glasses-basierter Dienstleistungssysteme, obwohl vor allem in der Logistik ein Einsatzgebiet und weiterer Forschungsbedarf in praxisorientierter Fachliteratur gesehen wird. Daher wurden eine explorative Fallstudie in Kombination mit einer systematischen Literaturstudie durchgeführt, um Use Cases für Smart Glasses in der Logistik zu identifizieren und aufzeigen, welche Use Cases priorisiert betrachtet werden sollten sowie welche Implikationen für zukünftige Forschung bestehen. In der Literatur wurden bisher vor allem Use Cases im Bereich Kommissionierung und Wartung bspw. der Stapler gesehen. Die Fallstudie eröffnete weitere Use Cases wie die Dokumentation von Beschädigungen oder die automatische Erfassung von Beständen (Inventur). Somit zeigte die explorative Fallstudie, dass Smart Glasses Einsatzpotenzial haben, um zukünftig Arbeitsprozesse zu digitalisieren. Forschungsbedarf und somit Aufgaben dieser Digitalisierung werden vor allem in Bezug auf die technische Machbarkeit und Integration in die Gesamtarchitektur gesehen, sowie im Bereich Datenschutz und organisatorischer Eingliederung bspw. durch Anpassung entsprechender Prozesse und des Datenmanagements. Forschungsbereiche der Wirtschaftsinformatik, die bei der Gestaltung des Einsatzes von Smart Glasses einbezogen werden sollten, sind somit u.a. Dienstleistungsmanagement und -engineering, Informationsprivatheit und Datenschutz aber auch Interface-Design, Unternehmensarchitekturmanagement und Geschäftsprozessmanagement. Somit leistet dieses Paper einen Beitrag zur Forschung im Bereich des ganzheitlichen Designs von Dienstleistungssystemen und zukünftiger Aufgaben digitaler Arbeit. Es liefert die Basis für weitere Forschung sowohl für (a) die gestaltungsorientierte Wirtschaftsinformatik, indem einzelne Szenarien mit Smart Glasses priorisiert zur Umsetzung vorgeschlagen werden und welche Herausforderungen dabei zu beachten sind, als auch (b) für die verhaltensorientierte Forschung, u.a. welche Bereiche priorisiert betroffen sein werden und welche Fragestellungen sich bspw. hinsichtlich Akzeptanz und neuer Arbeitsformen ergeben.

6 Danksagung

Dieser Beitrag ist Teil des Projekts *Glasshouse*, welches vom Bundesministerium für Bildung und Forschung (BMBF) unter dem FKZ. 01FJ15062 gefördert wird.

Literatur

1. Niemöller, C., Metzger, D., Fellmann, M., Özcan, D., Thomas, O.: Shaping the Future of Mobile Service Support Systems – Ex-Ante Evaluation of Smart Glasses in Technical Customer Service Processes. Informatik 2016, Klagenfurt (2016)
2. Rauschnabel, P.A., Ro, Y.K.: Augmented reality smart glasses: an investigation of technology acceptance drivers. Int. J. Technol. Mark. 11, 123–148 (2016)
3. Ernst, C.-P., Stock, B., dos Santos Ferreira, T.: The Usage of Augmented Reality Smartglasses: The Role of Perceived Substitutability. AMCIS 2016, San Diego (2016)
4. Hein, E.D.W., Rauschnabel, A.P.: Augmented Reality Smart Glasses and Knowledge Management: A Conceptual Framework for Enterprise Social Networks. In: Rossmann, A., Stei, G., and Besch, M. (Hrsg.) Enterprise Social Networks: Erfolgsfaktoren für die Einführung und Nutzung – Grundlagen, Praxislösungen, Fallbeispiele. 83–109. Springer Fachmedien, Wiesbaden (2016)
5. Niemöller, C., Metzger, D., Thomas, O., Ickerott, I., Till, S., Mollen, T., Neumann, T., Hucke, S.: Smart Glasses zur Unterstützung von Logistikdienstleistungen – Bedarfsorientierte Informationsbereitstellung zur Prozesssteuerung. Productivity (2015)
6. Brandl, P., Michalczuk, R., Stelzer, P., Bergles, K., Poggenburg, J., Sandtner, K.: Assist 4.0: Datenbrillen-Assistenzsysteme im Praxiseinsatz. Mensch & Computer 2014. 259–264 (2014)
7. Bitzen, S., Buttgereit, D., Hünefeld, R.: „Augmented Reality": Mehrwert oder Spielerei? Aachener Kolloquium für Instandhaltung, Diagnose, Anlagenüberwachung. 389-399 (2012)
8. ÇİÇEK, M.: Wearable technologies and its future applications. Int. J. Electr. Electron. Data Commun. 3, 45–50 (2015)
9. Yin, R.K.: Case Study Research: Design and Methods. Sage, London (2009)
10. Myers, M.: Qualitative Research in Business & Management. Sage, London (2009)
11. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Q. 26, xiii–xxiii (2002)
12. Fettke, P.: State-of-the-Art des State-of-the-Art. Wirtschaftsinformatik. 48, 257–266 (2006)
13. Verband der Hochschullehrer für Betriebswirtschaft e.V.: VHB Teilrating Logistik.
14. Bogner, A., Menz, W., Littig, B.: Das Experteninterview – Theorie, Methode, Anwendung. VS Verlag für Sozialwissenschaften (2009)

15. Bürki, R.: Klimaänderung und Anpassungsprozesse im Wintertourismus, Ostschweizerische Geographische Gesellschaft. 6 (2000)
16. Misoch, S.: Qualitative Interviews. De Gruyter Oldenbourg, Berlin (2015)
17. Sonnenberg, C., vom Brocke, J.: Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research. DESRIST 2012. 381–397 (2012)
18. Mayring, P.: Einführung in die qualitative Sozialforschung: Eine Anleitung zu qualitativem Denken. Beltz, Weinheim (2002)
19. Oates, B.J.: Researching Information Systems and Computing. Sage, London (2006)
20. Schumann, S.: Repräsentative Umfrage: Praxisorientierte Einführung in empirische Methoden und statistische Analyseverfahren. Oldenbourg Verlag, München (2012)
21. Chen, C.T., Cheng, H.L.: A comprehensive model for selecting information system project under fuzzy environment. Int. J. Proj. Manag. 27, 389–399 (2009)
22. Kühnapfel, J.: Nutzwertanalysen in Marketing und Vertrieb. Springer-Verlag (2014)
23. Jain, A.K.: Data clustering: 50 years beyond K-means. Pattern Recognit. Lett. 31, 651–666 (2010)
24. Caliński, T., Harabasz, J.: A dendrite method for cluster analysis. Commun. Stat. 3, 1–27 (1974)
25. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Q. 28, 75–105 (2004)
26. Walls, J.G., Widmeyer, G.R., El Sawy, O. a: Assessing Information System Design Theory in Perspective: How Useful was Our 1992 Initial Rendition. J. Inf. Technol. Theory Appl. 6, 43–58 (2004)
27. Dibia, V.: An Affective, Normative and Functional Approach to Designing User Experiences for Wearables. SSRN Electron. J. 1–12 (2015)
28. Gandy, M., Ross, D., Starner, T.E.: Universal design: Lessons for wearable computing. IEEE Pervasive Comput. 2, 19–23 (2003)
29. Smailagic, A., Siewiorek, D.: Application design for wearable and context-aware computers. IEEE Pervasive Comput. 1, 20–29 (2002)
30. Däuble, G., Özcan, D., Niemöller, C., Fellmann, M., Nüttgens, M.: Design of User-Oriented Mobile Service Support Systems – Analyzing the Eligibility of a Use Case Catalog to Guide System Development. Wirtschaftsinformatik (WI 2015). 149-163, Osnabrück (2015)
31. Rossi, M., Tuunainen, K.V., Pesonen, M.: Mobile technology in field customer service. Bus. Process Manag. J. 13, 853–865 (2007)
32. Breitschwerdt, R.: Informationstechnische Unterstützung mobiler Dienstleister: Eine Analogiekontruktion in der ambulanten Gesundheitsversorgung (Dissertation) (2013)
33. Bélanger, F., Crossler, R.E.: Privacy in the digital age: A review of information privacy research in information systems. MIS Q. 35, 1017–1041 (2011)
34. Chun, H., Lee, H., Kim, D.: The Integrated Model of Smartphone Adoption: Hedonic and Utilitarian Value Perceptions of Smartphones Among Korean College Students. Cyberpsychology, Behav. Soc. Netw. 15, 473–479 (2012)
35. Kim, K.J., Shin, D.-H.: An acceptance model for smart watches: Implications for the adoption of future wearable technology. Internet Res. 25, 527–541 (2015)
36. Merkel, T. Spitzhirn, M. Bullinger, A.C.: Einsatzszenarien für Smartphone und Wearables zur Verbesserung von Ergonomie, Arbeits- und Gesundheitsschutz. In: Bullinger, A.C. (ed.) Mensch 2020 – transdisziplinäre Perspektiven, Innteract. 107–117 (2015)
37. Venable, J., Pries-Heje, J., Baskerville, R.L.: A Comprehensive Framework for Evaluation in Design Science Research. DESRIST 2012. 423–438 (2012)
38. Sein, M.K., Henfridsson, O., Rossi, M., Lindgren, R.: Action Design Research. MIS Q. 35, 37–56 (2011)

Measuring the Agility of the IT Application Systems Landscape

Volker Nissen[1], Alexander von Rennenkampff[1]

[1] TU Ilmenau, FG WI für Dienstleistungen, 98684 Ilmenau, Deutschland
volker.nissen@tu-ilmenau.de

Abstract. A company's ability to change increasingly depends on the ability to change its IT, something referred to as "IT-agility" here. High IT-agility can contribute to increased business agility and thus create a competitive advantage. In this paper we focus on the IT application systems landscape, a resource of significant importance for the IT-agility and competitiveness of a company. To manage IT-agility it must be measurable. In our research, a goal hierarchy and a derived performance measurement (key figure) system was developed to measure and actively manage the agility of IT application systems landscapes. This measurement model is scalable from the measurement of individual domains to the entire IT application landscape. It has demonstrated its practicality in the context of several case studies.

Keywords: IT-Agility, IT Application Systems Landscape, Competitive Advantage, IT Controlling, IT Management.

1 Introduction

The business of companies and thus their business processes and products are changing over time. These changes almost always have an impact on the company's IT in the sense that IT systems need to be adapted. Surveys in recent years among IT managers as well as scientific contributions show that a key requirement for the IT organization is the ability to adapt to the needs of the professional business [5] [6].

Moreover, IT penetration of the core business processes in companies in recent decades has increased continuously. This will be even truer with digital transformation of industries gaining pace. Thus, the change ability of enterprises increasingly depends on the ability to change the IT [4] [21]. In business, we see that some companies are dealing better than others with the continuous need of change in their information systems (IS). Hence, we would like to understand, and ultimately measure and manage this ability of IT. To come up with a sound suggestion for these issues that we will refer to under the label of IT-agility is the central theme of our paper.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Nissen, V.; von Rennenkampff, A. (2017): Measuring the Agility of the IT Application Systems Landscape, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 425-438

2 Related Work – Literature Review

Based on a very recent and detailed literature review by Termer [33, p.24-30] the use of the “agility” term in information systems research will now be highlighted. By citation analysis, Termer identifies five major sources [p.24-25], amongst 57 relevant contributions, which dominate the definition of the term agility in the actual research literature: [9] [12] [25] [30] [22]. He also reveals that indeed a manifold of sources are used, so that overall a fragmented definition is apparent. It should be noted that in this analysis contributions dealing with agility in the limited context of software development are excluded. Also, papers using the term agility without defining it were excluded from further analysis. Moreover, research on the term “flexibility” is viewed as related, but not identical to “agility”.

Agility is understood by Sambamurthy et al. as an ability to recognize and use market potential rapidly and unexpectedly for competitors [30, p. 245]. This ability has two components: to probe the market (exploration) and to exploit the market (market arbitrage). Furthermore, agility is divided into three dimensions: customer agility, partner agility and operational agility. Even though IT in all three areas is assigned a supportive role, it is not part of the agility consideration [30, p. 246]. IT and the competence of using it are seen as the initiator and enabler for corporate actions, which are conducted in the three mentioned agility domains.

According to Overby [25, p. 121], who also sees IT in an enabling role for business, the capability of (enterprise) agility results in two significant basic skills: on the one hand the ability to notice changes in the corporate environment (sensing), and on the other hand the ability to conduct certain actions in a suitable way regarding these changes (responding). The sensing skill is here explicitly assigned to the tasks of anticipation and the prediction of environmental changes. The ability of responding includes moreover the essential pace to conduct the chosen action [16, p.127-128].

The definition of agility according to Goldman et al. [12] results from a production theoretic view. The term agile describes the ability of producers to be successful in changing, fragmented markets. Some constitutive features are specified: agility is dynamic and unlimited, context–specific, agile companies offensively embrace changes, and agility is offensive, as it creates chances for profit and growth. Termer argues [33, p. 31] that these characteristics overlap with features of flexibility, as described in the research literature. Flexibility refers to the capability of a corporation to be able to move rapidly from one known task to the next one. Here it is important that the appropriate situation is known beforehand and therefore the corresponding tasks to overcome this situation should also be defined in advance [12]. A corresponding questionnaire for the capturing of agility proves to be imprecise and subjective [16, p.127-128].

In the definition according to Dove [9] the ability of responding is especially emphasized: “We look at agility as deriving from both the physical ability to act (response ability) and the intellectual ability to find appropriate things to act on (knowledge management). Agility is expressed as the ability to manage and apply knowledge effectively so that an organization has the potential to thrive in a continuously changing and unpredictable business environment.” [9, p.9]. Hence two essential features of agility are emphasized: on the one hand the ability to act and on the other hand the ability to identify the right things for acting on. Basically the tasks of knowledge management are linked therewith [9, p.9-16]. Products, processes, practices and people are the primary action fields for agility [9, p.83, 163-180]. Dove reveals basic approaches, which contribute to better response ability. These include the abilities to reuse, reconfigure and rescale, which are used for different production scenarios [9, p. 33, 38, 42 and 61], as well as the standardization of interfaces to enable the use of a loosely coupled resource pool [9, p.41].

While considering agility, the aspects of proactive [9, p. 92-99] and reactive changes are incorporated [9, p. 100-107]. Proactive changes are made possible by means of innovation or by taking up the leading position in a division. Its trigger is therefore within the company. Dove refers to reactive changes as the foundation of survivability of a company and as the basis for opportunistic behavior. Reactive actions are always evoked by (external) events, which require such a reaction.

In the definition of agility according to Seo and La Paz the ability to notice signals from the internal and external environment (perception) and the adequate answers (responding) to these signals are placed into the center of the considerations [32, p. 136]. In this process, an organization is subjected to a perpetually changing cycle to which it adapts better and better, and thus ensures survival in a dynamic environment. Possible negative aspects in dealing with agility are predominantly discussed in this literature source [32, p.137–138). For instance, correct information from the environment has to be admitted, which furthermore needs to be made available for an adequate evaluation, so that the decision makers can make a decision on the basis of relevant information. Likewise it must be possible that actions can be implemented in a reasonable timeframe. Information systems have to be capable of this speed without claiming a high management effort. These negative aspects can be mitigated by standardization and outsourcing as well as by a suitable organization and culture.

Termer [33, p.28-30] concludes that none of the references identified in his extensive literature review actually designed a tool that would help companies identify the level of IT-agility attained, or follow the effectiveness of measures to increase IT-agility over time. The only exception, according to Termer, are the works of the research group of Nissen [22] [23] [24] [29], where a key figure system for the measurement of IT-agility was developed and tested in several practical case studies. The present article actually continues the development of this key figure system.

Termer, in his PhD-thesis [33] defines IT-agility from a behavioral point of view, stating that “agile behavior, also agility, is understood to mean the self-motivated, purpose-oriented activities of an organization which aims at the aggressive design of the environment” [33, p.46, translated]. IT-agility here is influenced by three main aspects of company IT: the technical infrastructure, the IT-staff, and the organizational structure, including IT-processes [33, p.49].

For the practical analysis of IT-agility, Termer recommends, drawing a parallel to air combat maneuvers, “to define appropriate actions (maneuvers) and to keep corresponding expectations of their implementation. The concrete execution of the actions could be documented and evaluated (…).” [p. 225, translated] While we view this behavioral definition of IT-agility as interesting from a scientific perspective, it must be said that this understanding is not very helpful from a practical point of view, as we cannot define maneuvers with a company IT in real life. Therefore, our own considerations of IT-agility, which are given in the next section, draw on the abilities and characteristics of the IT to define, measure, and manage the level of IT-agility.

Finally, we want to add the definition of IT-agility as suggested by the influential paper of Conboy [8, p.340] in the context of information systems development (ISD), which is a more limited scope than the one we address in our paper. Conboy states that agility is "the continual readiness of an ISD method to rapidly or inherently create change, proactively or reactively embrace change, and learn from change while contributing to perceived customer value (economy, quality, and simplicity), through its collective components and relationships with its environment." As one can see, there is some similarity to components in the previous definitions, particularly by Dove [9] and Termer [33]. Agility is seen as having reactive as well as proactive parts, ultimately thriving for customer value. We shall use these common aspects in our own conceptualization of IT-agility.

3 Conceptualization of IT-Agility

Our conceptualization of IT-agility is based on the underlying theoretical rationale of the resource-based view of strategic management (RBV). The RBV [1] [27] [37] places the heterogeneous equipment of companies with internal resources as a source of competitive advantage in the center. It is argued that not the sole possession of these resources is the cause of the success of a company, but additionally appropriate employee skills and management skills are needed who know to take advantage of the potential of resources. This will be reflected in our three pillars of IT-agility.

For the characterization of (strategically relevant) resources, various properties were defined in the academic discourse. According to the VRIS-framework of Barney, a resource is valuable, rare (or even unique), inimitable and non-substitutable, i.e. cannot be replaced by other equivalents [1, p 105]. Other authors have varied these properties and added in particular the usability and immobility aspects [35].

Applying the principles of RBV on the subject field of IT, the mere possession of IT does not lead to competitive advantage. Carr has described this situation very strikingly, in which he referred to IT as a commodity without strategic relevance [7]. This can be explained by his narrow viewing angle to pure technique. But strategic value of IT can be achieved through its effective and efficient use. The value contribution of IT to business then becomes a multi-facetted concept, where IT-agility is one component, next to others such as service orientation or enabling compliance [34]. Moreover, the target role of IT as an innovation engine for the company has been stressed in publications recently [17], fueling the understanding that IT is not a mere reactive enabler, but should also have a proactive function in the company.

In our conceptualization, a high IT-agility can be a strategic asset for companies, particularly in turbulent environments, and with many IT-based products or processes. Overall IT-agility is the result of agility in its three constituting parts ("pillars of IT-agility"), which we draw from the RBV and previous research in [22] [23] [24] [29]

[33]: 1. IT-infrastructure, in particular the architecture of the IT application systems landscape, 2. IT-organization / IT-processes, and 3. IT-staff / IT-management. This contribution deals only with the first of these three pillars.

Looking at the IT infrastructure of enterprises, it can be said that not all of its components are equally sources of competitive advantage. In particular, hardware and standard software are not strategically relevant. However, if the entire IT application systems landscape is considered, the criteria of the resource-based approach can be met, as was shown in [24]. Thus it can be stated that with the provision of an agile IT application systems landscape a strategic resource can be created that allows for a sustainable competitive advantage.

The elements of the IT application systems landscape are the application systems (applications and associated data), their interfaces as well as the domains and functions of a company. The architecture of the IT application systems landscape (short: application landscape) describes the application systems, their relationships and structure based on business-related criteria (domains and functions). The (nonexistent) IT architecture of obsolete application landscapes prevent easy maintenance and results in a low ability to change. This is supported by many literature sources, e.g. [11] [19] [28], which view IT application landscapes that have over many years grown uncontrolled as a main cause of the lack of IT-agility in companies. Moreover, there is a common believe that the architecture of the IT application systems landscape represents in fact a key differentiator between highly agile and less agile companies [10] [31] [23].

When dealing with change two archetypal forms can be distinguished: reactive and proactive. Many contributions in the IS literature define a passive coping with change as reactive, whereas an active, internally driven change intention is defined as proactive. When it comes to the type of change required in IT, one can basically distinguish capacitive and functional aspects. Capacitive requirements relate to scalability, i.e. the ability of IT to respond to growing business volumes, or performance, i.e. provide a constant response time even with changing demand volumes. Functional requirements relate to features, products, and processes of IT.

In the literature on agility there are some components of a definition that frequently appear: strategic thinking, identification of opportunities that create business value, proactive (innovative) and reactive (quick response) components, embracing change. We include the aforementioned considerations in our definition of IT-agility:

IT-agility is the ability of a company's IT to respond very quickly (preferably in real time) to changing capacity demands and changing functional requirements, as well as to use the potential of information technology in such a way that the business scope of action of the company is extended or even redesigned.

To manage IT-agility it must be measurable. Today no method is available that measures IT-agility and especially the agility of the IT architecture based on observable factors. Based on this situation, the following research question is derived:

How can the agility of the IT application systems landscape, as an important part of the IT-agility of a company, be measured and actively managed, based on objectively observable characteristics?

In the following, the development of a corresponding measurement instrument is outlined.

4 Method

4.1 Design Science Approach

The design science research (DSR) approach as a methodological framework seems well suited to answer the research question described. On the one hand the lack of measures of IT-agility and, more specifically, the agility of IT application systems landscapes is a relevant issue that comes from business practice. On the other hand, the development of a key figure system for measuring the agility of application landscapes is a design activity [15]. The specific sequence of the research activities in our investigation follows the popular DSR Methodology Process Model of Peffers et al [26]. From a scientific perspective, our contribution is the developed hierarchy of goals (and associated key figures) as a comprehensive model to explain the relationships between architecture principles and the agility of IT systems landscapes. The corresponding key figure system is the tool for measuring and actively managing IT-agility. In the following, we describe essential components and results of the iterative design process, which is grounded in a structured literature review, expert interviews and case studies. We report on results from the last design cycle.

4.2 Development of a Goal Hierarchy

First, the relevant parameters were determined for the design of the performance measurement system, i.e. those properties of IT application system landscapes that have an impact on the IT-agility. This was done through a structured literature review following Webster and Watson [36], not detailed here for reasons of space. The full description can be found in [29]. It revealed 29 relevant papers from which design input could be derived for measuring the agility of IT application systems landscapes, describing objectives in the context of IT flexibility and IT-agility, to be achieved with suitable measures. The goal network in Figure 1 shows the goals that were repeatedly mentioned. Bold relations are dominant. These are discussed frequently in the literature and have been further confirmed in a series of semi-structured interviews with 18 experts in the field. Suggestions obtained therefrom were incorporated into the final design of the goal hierarchy and associated indicators. In the semi-structured expert-interviews we followed [3] and questioned both IT top managers as well as IT architects. Moreover, in order to complete the picture, and also to get an external IT view, managers from the business were included that have a strong link to IT. Finally, some IT-related business consultants were also interviewed.

Along the dominant goal relations a goal hierarchy is constructed starting from the top goal. The goals connected by relations are adopted in the goal hierarchy. In addition, goals are further sharpened or detailed (Figure 2). Very few elements from the literature have not been incorporated, since they do not satisfy one or more of the required quality criteria. An example is the goal of high parameterization of the application landscape, meaning the ability of an application system to implement a business change without programming. For this goal the necessary data can in practice not be collected with reasonable effort, as was revealed in both the expert interviews and several practical case studies [24] [29].

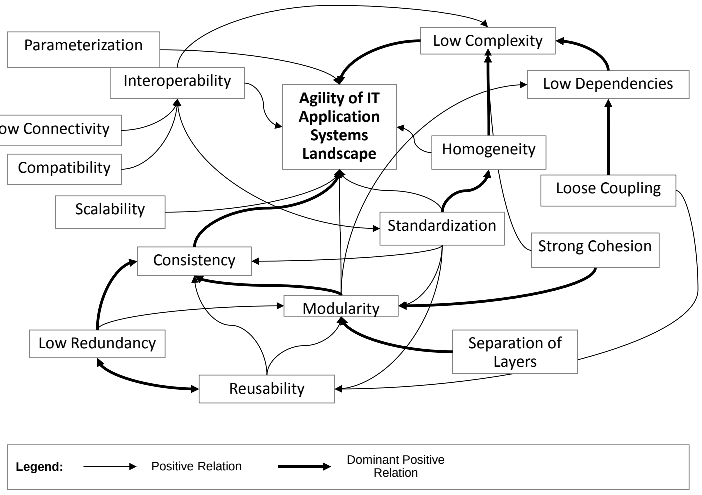


Figure 1. Goal relations from the analyzed literature

The goal hierarchy consists of five levels. On the first level is the overall goal "high agility of the IT application systems landscape." This is divided into the goals "high functional agility" and "high capacitive agility". The process is continued to the fifth level containing nine elementary goals, briefly characterized below.

Low Connectivity: The entire IT application systems landscape is regarded as a network of connected application systems. Low connectivity requires that in an application landscape as few application systems are interdependent as sensible.

Adequate Coupling: Adequate coupling requires that domains are internally closely and externally loosely coupled.

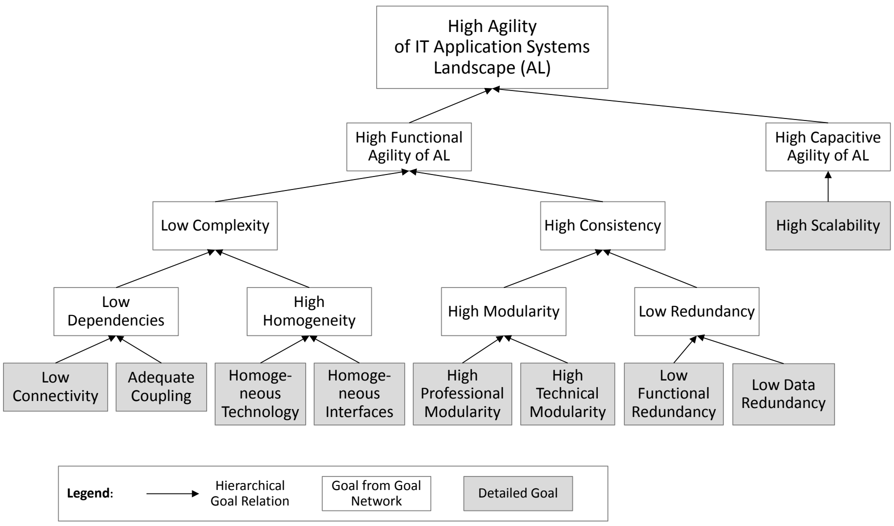


Figure 2. Goal hierarchy for the agility of an IT application systems landscape

Homogeneous Technology: The uniform application of few, dominant technologies facilitates the portability and compatibility of application systems. Functions can be easily re-used, which enables changes and contributes to IT-agility.

Homogeneous Interfaces: Uniform interface technologies reduce complexity. Data can easily be exchanged between applications.

High Professional Modularity: The structure of the application landscape of the company should be based on the company's business (processes and organization). The more two business processes, business functions or departments are linked together, the closer should be coupled the supporting application systems.

High Technical Modularity: Each application system should have a clear unambiguous technical purpose. Engels et al. differentiate four categories [11]: interaction components, process components, function components, inventory. High technical modularity requires that each system of the IT application landscape can be assigned to precisely one of the four established software categories.

Low Functional Redundancy: Ideally, a professional function is implemented only at a single point in the application landscape. If this function is used elsewhere, the existing implementation should be re-used. Multiple efforts for the implementation and for ongoing maintenance can thus be avoided.

Low Data Redundancy: Databases should be managed each by a single application system and all applications that require these data need to access the respective application system. Low redundancy of data is only required at the logical level.

High Scalability: Both the ability of the application system for parallel execution of processing steps as well as the capability of the technical infrastructure to scale horizontally should be evaluated jointly. The structure of the application landscape has only a supportive indirect influence on scalability. The direct factors affecting scalability can be found in the architecture of individual applications as well as in the nature of the technical infrastructure.

5 Artifact Description

5.1 Key Figure System

In order to manage the development of an IT application systems landscape towards a high agility, the attainment of the goals set out must be made measurable. To this end, we follow the approach suggested by Kütz [18]. For the documentation key figures profiles are used. Table 1 gives an example. The full key figure system is depicted in Figure 3.

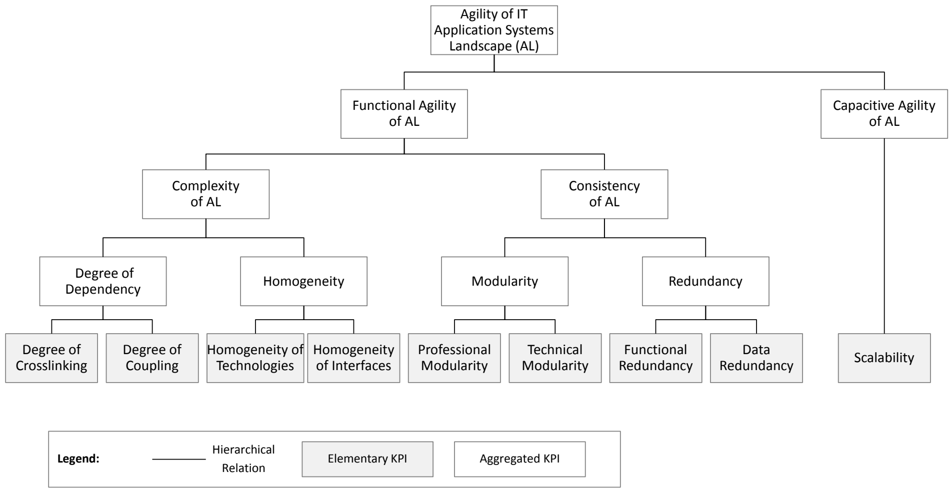


Figure 3. Key figure system to measure the agility of an IT application systems landscape

Table 1. Example of a key figure profile

KPI	
Technical Modularity ($MODU_{TECH}$)	
Supported goal	References
High technical modularity	[14, 15]
Question	
How well can the application systems of the IT application systems landscape be assigned to a unique technical software category?	
Description	
Technical modularity measures how well the application systems can be assigned to software categories. Software categories characterize different tasks (e.g. interaction, function or inventory) of application systems. Each application system should, ideally, be assigned a unique software category.The technical modularity sets the number of software category overruns of application systems in relation to the number of all application systems of the application landscape. The fewer overruns exist, the higher the value of the measure.	
Calculation	Standardization
$MODU_{TECH} = \frac{\sum(\mathrm{SK(a)} - 1)}{\mathrm{ELEM(AL)}}$	5 – $\mathrm{MODU_{TECH}} < 0.2$ 4 – $0.25 > \mathrm{MODU_{TECH}} \geq 0.2$ 3 – $0.35 > \mathrm{MODU_{TECH}} \geq 0.25$ 2 – $0.5 > \mathrm{MODU_{TECH}} \geq 0.35$ 1 – $MODU_{\mathrm{TECH}} \geq 0.5$
Data	
SK(a) = number of software categories application system „a“ can be assigned to ELEM(AL) = number of application systems in the entire application landscape	

A company needs to set target values, in accordance with the importance of IT-agility and after consideration of other strategic goals, and track the achievement of these target values. The performance measurement system can be used not only for measuring entire application landscapes, but also for parts of it. For clarity, the figures are, however, always defined for the whole IT application systems landscape. If, for example, the KPIs should be used for the measurement of individual domains, the concept of the application landscape must be replaced with the appropriate domain name. The potential user base of the performance measurement system is the top leadership circle of the IT organization as well as the enterprise architects of a company.

Below the elementary KPIs are briefly described. The indicators and thresholds in standardization were presented as part of the iterative design process in the expert interviews and used in several case studies, and have been intensively discussed.

Degree of Crosslinking: With the degree of crosslinking, the degree of internal dependencies of the application landscape under consideration is measured. The actually existing links between the application systems are set in proportion to the maximum possible number of links. The degree of crosslinking measures the complexity of the application landscape based on the connections between the application systems. The more application systems are connected together, the higher are the dependencies, and thus also the complexity of the application environment.

Degree of Coupling: The application systems or sub-domains contained in a domain should be more closely linked than the domains are connected to each other. Interfaces of a domain can be measured by the "coupling ratio", weighted by the number of application systems. The coupling ratio for a domain determines the ratio between the number of interfaces within a domain and the number of application systems contained therein. The average coupling ratio of the domains is assessed in relation to the coupling ratio of the entire application landscape. The greater this ratio, the stronger the domains are linked internally and the higher the degree of coupling.

Homogeneity of Technologies (Applications): The homogeneity of application systems technology measures the complexity that arises because different technologies are used in parallel in the application landscape. In addition to the number of different technologies also their distribution in the IT application systems landscape is measured. A high number of used technologies results in low technology homogeneity and thus leads to a high complexity of the application landscape.

Homogeneity of Technologies (Interfaces): The technology homogeneity of interfaces measures the complexity that arises when different interface technologies are used in parallel. In addition to the number of different interface technologies also their distribution in the application landscape is measured.

Professional Modularity: The professional modularity is the measure of the unambiguous assignment of application systems to domains. It sets the number of domain overruns of application systems in relation to the number of all application systems of the application landscape. The fewer domain overruns, the higher the value.

Technical Modularity: for details see Table 1

Functional Redundancy: The functional redundancy measures how many business functions have been implemented several times (and to what extent) in different applications. For this, the sum of all redundancies of functions in the application landscape is set in proportion to the total number of IT-supported functions.

Data Redundancy: The data redundancy measures how many data are maintained multiple times in different application systems and to what extent. For this, the sum of all redundancies of data stores in the IT application systems landscape is assessed in relation to the total number of data stores.

Scalability: This KPI measures the proportion of the application landscape, which is designed scalable. For this purpose, the proportion of the application systems is measured, where scalability was explicitly specified as a non-functional requirement in the implementation. This is then multiplied by the ratio of scalable infrastructure (hardware) components. The interviewed experts considered this KPI useful, but noted that the necessary data are rarely available in the required quality and completeness in practice. This has also been confirmed in the case studies conducted.

5.2 Aggregation of Key Figure Values

The hierarchical aggregation runs along the described goal hierarchy. Since the assessment of all key figures is carried out in an identical manner and standardized to the value range [1 ... 5], the calculation is made simple. The lowest level consists of the already explained elementary KPIs. The aggregate indicators in the levels above are always put together by the same aggregation rule, a weighted additive aggregation of subordinated key figures.

The weighting factors of the indicators have values between 0 and 1, and their sum in an aggregation is always equal to 1. As a result, the lower-level KPIs can be weighted relative to one another, and the result is always on a scale between the values 1 and 5. Through the weighting factors, companies can set priorities, if they consider parts of the goal hierarchy to be particularly desirable. For a benchmarking of various companies or parts of the same company, however, the weighting factors must be kept constant, otherwise the comparability of the results is not given

6 Discussion

The motivation of our research and the problem to be solved is the lack of instruments to measure and actively manage the IT-agility (not only) in the field of the IT application systems landscape. This problem is of considerable practical importance across industries, since the IT-agility can be seen as an integral part of the value contribution of IT in an enterprise [34]. Our research makes contributions on the scientific as well as the practical side of this problem.

From a scientific point of view, our contribution is the developed hierarchy of goals (and associated key figures) While we acknowledge that agility as a research topic has been treated in a sizable number of contributions in the literature, the agility of the IT architecture is rarely discussed [14]. To our knowledge, we provide the first comprehensive model to explain the relationships between architecture principles and the agility of IT application systems landscapes.

The applicability and usefulness of the key figure system was tested in several case studies in practice. Here, the approach followed recommendations of Yin [38] and Benbasat et al. [2]. In selecting the companies studied emphasis was placed on finding relatively different companies in order to examine a wide range of possible scenarios for the key figure system. For reasons of space, the case studies cannot be presented here. However, details can be found in [29]. In order to ensure the validity and reliability, different data sources were used in the companies, such as architecture data bases, architecture graphs, architecture concepts and interviews with enterprise architects. In addition, a database has been set up for each case study that includes, separated from the raw data, a comprehensible analysis and calculation of key figures.

In our work, case studies and expert interviews were combined to evaluate the performance measurement system. This combination of evaluation methods is frequently used in application-oriented work when no similar model exists [20]. It turned out that through the developed indicators the agility of the IT application systems landscape can be measured and controlled over time, thus making a substantial practical contribution.

We now compare our approach with the DSR guidelines as provided by Hevner [15, p.83]. Guideline 1 (Design as Artifact): A key figure system is an artifact in the sense of Hevner. Guideline 2 (Problem Relevance): The measurement and management of IT-agility are issues of substantial relevance that add to the value contribution of IT in practice. Guideline 3 (Design Evaluation): Although, because of a lack of space, the evaluation had to be largely omitted from this paper, the goal hierarchy and associated key figure system have undergone substantial evaluation through expert interviews and practical case studies. Guideline 4 (Research Contribution): The key figure system and the goal hierarchy provide scientists with a new basis for the concept of IT-agility in the IT architecture of organizations. Guideline 5 (Research Rigor): We used a mixed-method approach combining a structured literature review with semi-structured expert interviews and practical case studies. All methods were designed after established methodology references as mentioned in the text. Guideline 6 (Design as a Search Process): Our research required several design cycles to come up with the final artifacts as presented here. The research on IT-agility in our group is an ongoing process since 2008. Guideline 7 (Communication of Research): This paper is our attempt to communicate the current results to the scientific community.

7 Conclusions

IT-agility is attributed a strong value proposition to corporate success, however, the required level of IT-agility is seldom achieved in companies. According to our expert interviews, the demand for IT-agility is considered particularly high in companies with a strong IT penetration, complex and diverse products with short product life cycles, a strong end customer orientation and dynamic competitive environments. Based on design principles and with appropriate measuring instruments the IT-agility can be improved in practice.

In our research, a goal hierarchy and a derived performance measurement (key figure) system was developed to measure and actively manage the agility of IT application systems landscapes. This measurement model is scalable from the measurement of individual domains to the entire IT application landscape.

The key figure system was developed with a focus on the part of the application landscape that supports the core business processes of a company. In particular in the field of Business Intelligence, which often accounts for a significant portion of IT budgets, questions concerning the applicability of the suggested KPIs to the field of analytical applications may be of interest. Also interesting is whether the performance measurement system can be applied to enterprise networks (value chains).

Methodically, it would be desirable to quantitatively evaluate the key figure system. For this purpose, however, a large number of participating companies is necessary. A large-scale quantitative survey would also enable the validation of the proposed standardization (normalization of indicator results).

Also, the question of the cost of the build-up of IT-agility is not yet considered. Only when this has been clarified, together with the demand question (How much IT-agility is actually needed?), a statement about the optimum level of IT-agility in a company can be made.

While in this contribution the IT architecture was the focus, appropriate considerations are also required with respect to the other pillars of IT-agility, namely the IT organization and IT processes as well as IT staff and management. This represents a current focus of our research with the ultimate goal to make IT-agility measurable and manageable in all fields of action.

References

1. Barney J. Firm Resources and Sustained Competitive Advantage. Journal of Management 17 (1), 99–120 (1991)
2. Benbasat, I., Goldstein, D.K., Mead, M.: The Case Research Strategy in Studies of Information Systems. MIS Quarterly 11(3), 369–386 (1987)
3. Bortz, J., Döring, N. Forschungsmethoden und Evaluation. 5.ed., Springer, Berlin (2015)
4. Byrd, T., Turner, D.: An Exploratory Examination of the Relationship between Flexible IT Infrastructure and Competitive Advantage. Information & Management 39 (1), 41-52 (2001)
5. Capgemini (ed.) Application Landscape Report 2011 Edition, München (2011)
6. Capgemini (ed.) IT-Trends 2012. Business-Alignment sichert die Zukunft. München (2012)
7. Carr, N.G.: IT Doesn‘t Matter. Harvard Business Review 2003/5, 41-51 (2003)
8. Conboy, K.: Agility from First Principles: Reconstructing the Concept of Agility in Information Systems Development. Information Systems Research (20) 3, 329-354
9. Dove, R.: Response Ability. The language, Structure, and Culture of the Agile Enterprise. John Wiley & Sons, New York (2001)
10. Durst, M.: Wertorientiertes Management von IT-Architekturen. Springer, Berlin (2007)
11. Engels, G., Hess, A., Humm, B., Juwig, O., Lohmann, M., Richter, J.P., Voss, M., Willkomm, J.: Quasar Enterprise. dpunkt, Heidelberg (2008)
12. Goldman, S.L., Nagel, R.N., Preiss, K (1995) Agile Competitors and Virtual Organizations. Strategies for Enriching the Customer. Van Nostrand Reinhold, New York (1995)
13. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly (37) 2, 337-355 (2013)
14. Gronau, N.: Wandlungsfähige Informationssystemarchitekturen. GITO, Berlin (2006)
15. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly. 2004 28 (1), 75–105 (2004)
16. Hoek van, R.I., Harrison, A., Christopher, M.: Measuring Agile Capabilities in the Supply Chain. Int. Journal of Operations & Production Management 21(1/2), 126–148 (2001)
17. Kießling, M.: IT-Innovationsmanagement. Cuvillier, Göttingen (2012).
18. Kütz, M.: Kennzahlen in der IT. 3ed. dpunkt, Heidelberg (2008)
19. Linthicum, D.S.: Enterprise Application Integration. Addison-Wesley, Boston (2000)

20. Maske, P.: Multiperspektivische Evaluation des integrierten, interdisziplinären Vorgehensmodells. In: Maske P (ed.) Mobile Applikationen. 2.ed. Gabler, Wiesbaden,797–939 (2012)
21. Melarkode, A., Fromm-Poulsen, M., Warnakulasuriya, S. Delivering Agility through IT. Business Strategy Review 15 (3), 45-50 (2004)
22. Nissen, V., Mladin, A.: Messung und Management von IT-Agilität. HMD Praxis der Wirtschaftsinformatik 269, 42–51 (2009)
23. Nissen, V, Rennenkampff, A von, Termer, F: IS Architecture Characteristics as a Measure of IT Agility. In: Proceedings of the 17th AMCIS, Detroit. AIS, Paper 89 (2011)
24. Nissen, V., v. Rennenkampff, A.: IT-Agilität als strategische Ressource im Wettbewerb. In: Lang, M. (ed.): CIO-Handbuch. Vol. 2. Symposion, Düsseldorf, 57–90 (2013)
25. Overby, E., Bharadwaj, A.S., Sambamurthy, V. (2006) Enterprise Agility and the Enabling Role of Information Technology. Europ. Jour of Inform. Systems 15(2), 120–131 (2006)
26. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.A: Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24 (3), 45–78 (2008)
27. Penrose, E.T.: The Theory of the Growth of the Firm.Wiley, New York (1959)
28. Reddy, S.B., Reddy, R.: Competitive Agility and the Challenge of Legacy Information Systems. Industrial Management & Data Systems 102 (1), 5–16 (2002)
29. Rennenkampff, A.v.: Management von IT-Agilität. PhD Diss., TU Ilmenau, Ilmedia (2015)
30. Sambamurthy, V., Bharadwaj, A., Grover, V.: Shaping Agility through Digital Options: Reconceptualization of the Role of Information Technology in Contemporary Firms. MIS Quarterly 27 (2), 237–263 (2003)
31. Schelp, J., Winter, R.: Integration Management for Heterogeneous Information Systems. In: Desouza, K.C. (ed.): Agile Information Systems. Butterworth, Burlington 134–149 (2007)
32. Seo, D., La Paz, A.I.: Exploring the Dark Side of IS in Achieving Organizational Agility. Communications of the ACM 51(11), 136-139 (2008)
33. Termer, F.: Determinanten der IT-Agilität. Springer Gabler, Wiesbaden (2016)
34. Termer, F.; Nissen, V.; Dorn, F.: Grundlagen zur Konzeption einer multiperspektivischen Methode zur Messung des IT-Wertbeitrags. In: Proceedings of MKWI 2014, Paderborn, 2271-2283 (2014)
35. Wade, M., Hulland, J.: Review: The Resource-based View and Information Systems Research: Review, Extension, and Suggestions for Future Research. MIS Quarterly 28 (4), 107–142 (2004)
36. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26 (2), 13–23 (2002)
37. Wernerfelt, B.: A Resource-based View of the Firm. Strategic Management Journal 5 (2), 171–180 (1984)
38. Yin, R.K.: Case Study Research: Design and Methods. Sage Publications, New York (2009)

Analyzing the Trade-Off between Traditional and Agile Software Development – A Cost/Risk Perspective

Patrick Afflerbach[1], Simon Kratzer[1], Maximilian Röglinger[2], and Simon Stelzl[1]

[1] University of Augsburg, FIM Research Center, Augsburg, Germany
[2] University of Bayreuth, FIM Research Center, Bayreuth, Germany
{patrick.afflerbach, simon.kratzer, maximilian.roeglinger, simon.stelzl}@fim-rc.de

Abstract. Digitalization heralds a new era of enterprise IT. It challenges CIOs to find a balance between renovating legacy IT and seizing the opportunities of digital technologies to keep up with competitors and start-ups. This requires organizations to operate two software development modes simultaneously: the traditional and the agile mode. Despite substantial research on both modes, little is known about whether to implement distinct software development projects traditionally or agile. As a first step to addressing this gap, we propose a quantitative decision model that compares the cost and risk profiles of both modes associated with the implementation of a distinct project. The decision model integrates qualitative and quantitative characteristics of the project in focus and of the traditional and the agile mode. As for evaluation, we implemented the decision model as a software prototype and validated its behavior using sample projects as well as a sensitivity analysis.

Keywords: Software Development, IS/IT Management, Agile Methods, Decision Model.

1 Introduction

Digitalization is transforming many industries [1]. Digital technologies enable innovative business models and provide novel revenue opportunities. In their yearly CIO Survey, Gartner [2] relate digitalization to the third era of enterprise information technology (IT). While the first era refers to the initial implementation of corporate information systems, the second era focused on improving reliability and stability. Currently, new technological trends are shaping the third era in a way most CIOs are not prepared for [2]. There is a tension between "doing IT right […], doing IT safely […] and working the plan" and "doing IT fast, […] doing IT innovatively […] and adapting" [2].

Although the third era of enterprise IT requires different architectures, processes, and governance, it mostly impacts software development (SD) [1]. To tackle the challenges related to the tension mentioned above, researchers and industry experts suggest operating multiple SD modes within a company [1, 3]. The two most common modes are the traditional (TSD) and the agile SD mode (ASD) [2]. TSD is about upfront planning and maximizing control, ASD about risk-taking and dealing with uncertainty to

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Afflerbach, P.; Kratzer, S.; Röglinger, M.; Stelzl, S. (2017): Analyzing the Trade-Off between Traditional and Agile Software Development - A Cost/Risk Perspective, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 439-453

create innovation at a high pace [4]. Both research and industry call for operating TSD and ASD in parallel. Thus, there is a shift from talking about why and how to do agile toward finding an appropriate trade-off between TSD and ASD.

Most approaches that support deciding about the appropriate SD mode for a distinct SD project (SDP) rely on qualitative analysis [3, 5]. Although there is research on quantifying the characteristics of both modes, there are, to the best of our knowledge, no approaches that quantitatively analyze the decision between TSD and ASD with respect to the implementation of a distinct SDP [6]. Thus, we investigate the following research question: *How can organizations decide whether to implement a distinct SDP in TSD or ASD mode regarding a cost/risk perspective?*

As a first step to answering this question, we adopted the design science research (DSR) paradigm and propose a quantitative decision model, taking a cost/risk perspective [7]. Our model incorporates characteristics inherent to the SDP in focus as well as characteristics specific to TSD and ASD mode. Mapping project characteristics to the cost/risk profiles of both SD modes, our decision model addresses the TSD/ASD trade-off for a distinct SDP. Thereby, our model does not strive for exact estimations of an SDP's business value. Rather, it aims to compare the economic effects associated with both SD modes. We thus take a tactical perspective on the TSD/ASD trade-off instead of a budget planning and cost estimation perspective.

Following the DSR reference process [8], we first provide theoretical background on both SD modes as well as on cost, benefit, and risk management for SDPs. We also derive design objectives that characterize proper solutions to our research problem (*objectives of a solution*). We then introduce the design specification of our decision model (*design and development*), before we report on our evaluation activities so far (*evaluation*). We conclude by summarizing key results, reviewing our limitations, and suggesting topics for further steps to answer this research question.

2 Theoretical Background and Design Objectives

2.1 Traditional and Agile Software Development

SD involves activities grouped into four phases, i.e., specification, design, implementation, and testing [9]. Although all SD methods follow these phases, they differ in how and how often these phases are traversed. TSD mode is sequential or iterative, while ASD is evolutionary or iterative and incremental [10]. The philosophy of TSD is plan- and process-driven as well as formal and tool-oriented [11, 12]. Aitken and Ilango [3] list methods for TSD mode. Methods related to TSD mode rely on the upfront planning of requirements, documentation, and long development cycles [11, 13]. This makes TSD mode beneficial for projects whose scope and features are known upfront. Defects or changes in scope or requirements, however, result in expensive iterations of previous phases [9]. To outline the characteristics of TSD, we use the waterfall model, which sequentially traverses the four SD phases without any iteration or repetition [14].

Due to long development cycles and rising demand for more flexible methods, ASD mode evolved in the 1990s. In 2001, experienced software practitioners issued the Agile Manifesto [15]. The idea behind ASD is the prioritization of individuals and interactions over processes and tools, working software over comprehensive documentation, customer collaboration over contract negotiation, and responding to change over following a plan [15]. Dybå and Dingsøyr [13] as well as Lee and Xia [10] provide an overview of methods related to ASD mode. A well-known ASD method is SCRUM. The essential part of SCRUM is a sprint, an iteration that may last from one day to four weeks. After each sprint, there is a working software artifact ready to be released. ASD methods are primarily characterized as being able to react to changes in requirements and to encourage developers' creativity and autonomous work. As a disadvantage, ASD heavily relies on trustful relationships among customers, developers, and managers as well as a high amount of self-reliance of all participants. Furthermore, ASD often faces problems in a highly regulated and plan-driven environment [3]. Aitken and Ilango [3], Conboy [11], as well as Nerur et al. [12] compare TSD and ASD mode according to mode-specific characteristics.

The argumentation above revealed structural differences between TSD and ASD when implementing an SDP. These differences make the two modes differently suitable for specific projects. Thus, decisions about the appropriate SD mode must account for these differences. Against this background, we define the following design objective:

(DO.1) *Differences of TSD and ASD:* The decision on the appropriate SD mode must account for structural differences of TSD and ASD mode.

2.2 Costs and Risk in Software Development Projects

The dynamic environment in the third era of enterprise IT challenges managers to choose among different SDPs and to decide in which mode to implement these SDPs. In the last decade, researchers and practitioners proposed success criteria and frameworks for project valuation [16]. Techniques that help quantify an IS/IT project's value fit into five categories: net present value (NPV) methods, rate of return methods, ratio methods, payback methods, and accounting methods [17]. The cost-benefit analysis, a representative of the ratio methods, has evolved into the most popular approach to the valuation of IT projects [18]. With this popular approach, the qualitative and quantitative benefits of a distinct project are compared to implementation and operating costs. Finally, a project has a positive business value, if benefits exceed costs. Besides benefits and costs, risk evolved into a third recognized dimension of cost-benefit analysis [19]. As an SDP's largest benefits realize at run time of the related software product, they are out of our decision model's scope [20].

The estimation of an SDP's implementation costs is a well-covered topic in software engineering [6]. Thus, many software cost estimation models are available. These models can be grouped into five categories, i.e., model-based, expert-based, learning-oriented, regression-based, dynamics-based, and composite approaches. Model-based approaches propose algorithms that are specifically developed for the SDP context [9]. COCOMO is a model- and regression-based algorithmic approach for estimating the costs of projects conducted in TSD mode based on input parameters such as project size, scaling factors, and effort multipliers [21]. As a widely recognized approach to software cost estimation, we use the improved version COCOMO II in our decision model for projects executed in TSD mode. There are also cost estimation models for projects conducted in ASD mode. Cao et al. [22] suggest dynamics-based methods that can be adjusted to changing requirements. Yang et al. [23] as well as Benediktsson et al. [24] modify COCOMO II for ASD mode, adjusting the number and length of iterations as well as by incorporating additional efforts. As COCOMO II and the extensions proposed by Benediktsson et al. [24] use the same input parameters, their recommendations are comparable. Thus, we rely on these approaches to consistently compare the costs for implementing an SDP in TSD and ASD mode.

In general, risk is an uncertain event or condition that, if it occurs, has a positive or negative effect on a project objective [20]. In our study, we focus on IT project risks, as we investigate SDPs from a stand-alone perspective and focus on the implementation phase. Studies about IT project risk are provided by Schmidt et al. [25], Boehm [26], Bannerman [20]. Referring to these studies, the most important risks are the following ones: lack of top management support, misunderstanding of requirements, changes in requirements, lack of user involvement, and lack of adequate project staff. We aggregate these risks into three risk types that represent the major impacts of the above-mentioned risk factors on an SDP, i.e., the risk of requirement changes, the risk of delay, and the risk of defect. In our model, we adjust an SDP's implementation costs by risky costs originating from the risk types just introduced [27].

The preceding reasoning underscores the relevance of an integrated cost/risk-perspective when deciding about the appropriate SD mode for a distinct SDP. An integrated perspective is not only necessary as for economic factors, but also when it comes to the comparison of both SD modes. We define the following design objective:

(DO.2) *Integrated economic valuation:* The decision on the appropriate SD mode requires considering economic dimensions (e.g., costs and risks) in an integrated manner. This decision requires ensuring the comparability of both modes.

3 Design Specification

Our decision model systemizes and combines factors relevant for determining the appropriate SD mode for a distinct SDP. Thus, our unit of analysis is the implementation of a distinct SDP. In our basic setting, the management has already strategically decided to implement the SDP, i.e., the benefits outweigh costs and risk independent from the chosen SD mode. Only the decision about the appropriate SD mode is open. We assume that the organization has the capabilities and expertise to conduct the SDP in focus in both modes (Assumption 1). In our decision model, this operative decision is based on the costs and the risks associated with TSD and ASD mode. Although researchers point to qualitative factors beyond a cost/risk perspective, we focus on their economic impacts and model them implicitly via the costs and risks effects of the SDP in focus [28]. As stated, we abstract from benefits as they materialize at run time. For the same reason, we focus on those costs and risks that accrue during the SDP's implementation.

The decision about the appropriate SD mode involves two major challenges. First and inherent to almost all economic decisions, the decision requires resolving trade-offs. In our TSD/ASD trade-off, the advantages and drawbacks of TSD mode face those related to ASD mode. Second, the decision is multi-dimensional. In order to address the TSD/ASD trade-off, the involved decision-makers must integrate quantitative and qualitative factors. We address these challenges by integrating these factors and evaluating their effects economically from a cost/risk perspective.

3.1 General Setting

We first introduce the general setting of our model and outline the TSD/ASD trade-off. We also define the model's objective function. From an economic perspective, we can break down the TSD/ASD trade-off into differences regarding the cost/risk profiles of an SDP in both SD modes. There are projects for which the cost/risk profiles dovetail, with the TSD/ASD trade-off clearly tending to one side. We refer to such SDPs as TSD- or ASD-type projects. TSD is the preferable SD mode if the cost and risk advantages of TSD exceed those of ASD – and vice versa. There are also projects whose profile cannot be unambiguously assigned to either type.

To facilitate a structured presentation of the decision model, we first consider costs and risks separately, before integrating them. The deterministic base costs C are defined as those costs associated with the SDP's implementation. We model the base costs by using the cost estimation models COCOMO II by Boehm et al. [21] for TSD mode and the extensions of Benediktsson et al. [24] for ASD mode. Furthermore, we define the risk R as additional costs taking the risks of defects, requirement changes, and delays into account. Basically, there are two reasons why we chose COCOMO II for modeling base costs: First, COCOMO is popular and well-known among researchers and practitioners, which is why it increases the comprehensibility of our model [9]. Second, COCOMO II and the extension by Benediktsson et al. [24] ensure comparability as both methods use identical input parameters. In fact, our aim is to offer decision support on the comparison of TSD and ASD for a distinct SDP, and not to estimate costs as accurately as possible.

To evaluate the SDP's costs and risk depending on the chosen SD mode, we use the certainty equivalent method. The certainty equivalent method is an acknowledged approach for corporate decision-making in risky situations [29]. It has also proven to be useful for decision-making in IS/IT projects as well as in IS/IT project portfolio management [16, 30]. The certainty equivalent represents the certain amount of money that creates the same subjective utility for the involved decision-makers as the corresponding risky situation [27]. Applying the certainty equivalent enables comparing the deterministic base costs C and the risky additional costs R. For decision-makers with a constant absolute risk aversion α, measured via the absolute Arrow-Pratt measure, and for a normally distributed risk with expected value μ_R and variance σ_R^2, the certainty equivalent of a distinct SDP for a chosen SD mode is given by Eq. (1) [31]. As the certainty equivalent can be structurally decomposed into the expected costs and a risk premium, we refer to the certainty equivalent as risk-adjusted costs to outline both decisive components. On this foundation, we use the difference between the risk-adjusted costs that accrue when implementing the SDP in TSD or in ASD mode as our objective function, which is shown in Eq. (2).

$$\Phi = \mathbb{E}[-(C+R)] - \frac{\alpha}{2} \cdot VAR[-(C+R)] = -C - \mu_R - \frac{\alpha}{2} \cdot \sigma_R^2, \qquad \alpha \in \mathbb{R}^+ \tag{1}$$

$$\Delta\Phi = \Phi^{\text{TSD}} - \Phi^{\text{ASD}} \tag{2}$$

In case $\Delta\Phi > 0$, the SDP should be implemented in TSD mode, and vice versa. Below, we analyze the costs, their components, and the corresponding parameters. After that, we provide more details on the risk component.

3.2 Cost Perspective

First, we focus on the base costs of an SDP. The base costs for implementing an SDP depend on the chosen SD mode, because TSD and ASD mode feature different cost structures [3]. In TSD mode, the SDP is organized in a long phase which is processed sequentially in a single iteration. In contrast, ASD mode is organized and processed in several smaller iterations, so-called sprints. As mentioned above for modeling the base costs we use COCOMO II by Boehm et al. [21] for TSD mode and its extension from Benediktsson et al. [24] for ASD mode.

The base model of COCOMO II describes the SDP's effort in TSD mode in of person-months by taking the project size to the power of a scaling exponent E (Eq. 3). The number of person-months multiplied with the staff costs leads to the base costs. The variable size is measured in thousand source lines of code (KSLOC). Other measurement units for size (e.g., function points) are applicable as well. The constant A is an effort coefficient for productivity and the parameters $EM_1, \dots, EM_{16}$ are effort multipliers that represent factors like product complexity, programmer capability, or platform volatility. The factor a thus adjusts the effort regarding productivity. The scaling exponent, which indicates economies $(E < 1)$ or diseconomies of scale $(E > 1)$, comprise a constant B and scale factors $SF_1, \dots, SF_5$ accounting, for example, for development flexibility, team cohesion, and process maturity. If all scale factors are rated "very high" and B is calibrated as COCOMO II suggests, E equals at least 0.91, implying economies of scale. Economies of scale can hardly be realized in SDPs. In almost every possible parameterization of COCOMO II, the cost function exhibits an exponent between 1.0 and its maximum 1.226, indicating diseconomies of scale. This is as communication and integration overheads increase over-proportionally with the project size. Eq. (3) represents the base model for TSD mode Boehm et al. [21].

$$PM_{\text{TSD}} = a \cdot size^E \qquad \text{with } E = B + 0.01 \cdot \sum_{j=1}^{5} SF_j \text{ and } a = A \cdot \prod_{i=1}^{16} EM_i \tag{3}$$

Benediktsson et al. [24] adjust the base model of COCOMO II for ASD mode. They divide the project size into sprints, a transformation that causes an initial effort for defining the breakdown structure. Moreover, sprints are interpreted as sub-projects. Consequently, Benediktsson et al. [24] apply the exponent E to the sprint sizes and sum up the resulting efforts, arguing that smaller sprints require less overhead and planning compared to the single large phase in TSD mode. In reality, the overall project length may vary between the different SD modes. However, there is no evidence that one mode is dominating the other. Therefore, and for reasons of comparability, we assume the project length to be identical for TSD and ASD mode (Assumption 2). Our model setup would allow for length discounts in favor of ASD by subtracting a certain number of sprints from the length calculated for project implementation in TSD mode. Such a parameter, however, would be subject to extreme estimation inaccuracies and increase the complexity of our decision model. Furthermore, sprints can differ in size and content, which depends very much on the SDP at hand. We assume that all sprints take a constant amount of time (e.g. two weeks, four weeks) in case the SDP is implemented in ASD mode (Assumption 3). Thus, we divide the project length calculated for TSD mode into n equi-length sprints in ASD mode and apply the approach of Benediktsson et al. [24] [32]. Eq. (4) shows the base model for implementing the SDP in ASD mode as an adaption of the TSD mode estimation.

In Eq. (4), the parameter d_n models the overhead of the total effort required to define the breakdown structure. It depends on the total number of sprints [24]. Thereby, d_n is calibrated by a linear interpolation between the relative effort for defining one sprint e and for defining the largest number of sprints f (Eq. 5). The parameter c represents the overhead costs per sprint and accounts for the additional overhead within single sprints. The number of sprints n is gained from the calculation of the project length in COCOMO II multiplied with the number of sprints per month. The exponent F consists of similar components as its counterpart for TSD mode. The constants G (called C in COCOMO II) and D are obtained by calibrations based on former projects [21].

$$PM_{\text{ASD}} = d_n \cdot a \cdot size^E + n \cdot a \cdot \left(\frac{size}{n} \cdot (1 + c)\right)^E \tag{4}$$

with $n = G \cdot (PM_{\text{TSD}})^F \cdot sprints\ per\ month$ and $F = D + 0.02 \cdot 0.01 \cdot \sum_{j=1}^{5} SF_j$ and

$$d_n = (f - e) \cdot \frac{n-1}{n-N} + e \tag{5}$$

Figure 1 illustrates the cost/size relationships for projects implemented in TSD and ASD mode. Costs are measured in Monetary Units [MU], size in KSLOC. Since SDPs typically exhibit diseconomies of scale, we apply an exponent larger than 1 [21].

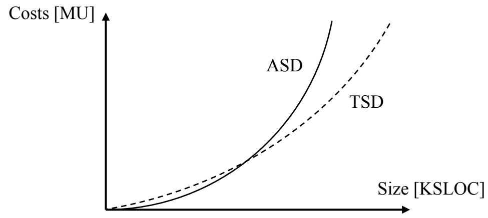


Figure 1. Costs of an SDP with diseconomies of scale depending on size

For $E \neq 1$, the two cost curves cross each other at a distinct intersection point. For small projects, ASD mode appears more favorable. This is as the initial effort required for creating a small number of sprints in ASD mode is small compared to the initial overhead for planning in TSD mode. Thus, TSD mode needs a minimum project size to justify the initial overhead. From a pure cost perspective, TSD mode is becoming more preferable with an increasing project size. This is as the initial overhead is distributed over the project length. For large projects, the initial overhead for ASD mode is becoming relatively expensive as planning and managing sprints is getting complex [33].

3.3 Risk Perspective

In our analysis, we focus on three risk types, i.e., the risk of defect, risk of requirement changes, and risk of delay. Defects, changes in requirements, and delays can happen several times during implementation of an SDP. We model each risk type as a normally distributed variable with expected additional costs and a variance per sprint (Assumption 4) [27, 30]. Typically, the expected value μ per risk type is denoted in percent of the corresponding total costs C depending on TSD or ASD mode. Furthermore, the variance σ^2 of each risk type is measured in percent of the expected value.

We add this risk perspective to our decision model in order to incorporate the specific characteristics of TSD and ASD mode. The risk of defect covers the risk of having major defects in the software product due to developer mistakes [25, 26]. Boehm and Turner [34] model this risk based on an SDP's criticality. A large project size combined with a lack of upfront planning and inexperienced developers, for example, can cause such defects. A defect typically results in negative economic effects, which we model by expected additional costs resulting from a defect μ_{DEF} and their variance σ^2_{DEF}. Since there is a detailed ex-ante planning in TSD mode, the expected additional costs and their variance are typically smaller as compared to ASD mode [12]. The risk of requirement changes, which Boehm and Turner [34] call the dynamism of a project, addresses change requests during project implementation. Due to regular releases, frequent stakeholder feedback as well as an environment geared toward refactoring and backlog refinement, the expected additional costs for requirement changes μ_{CHG} and their variance σ^2_{CHG} are generally lower in ASD mode than in TSD mode [10]. The risk of delay addresses the risk of missing a deadline or not being fast enough regarding time-to-market. This risk is particularly high if the software product exhibits a short time-to-market. Delays result in expected additional costs μ_{DEL} and variance σ^2_{DEL} from losing market share to competitors or from financial penalties due to missed deadlines. ASD mode can better mitigate the risk of delay as running software increments can be released more flexibly [5]. Thus, expected additional costs and the variance are generally higher in TSD mode.

In order to aggregate the three risk types, we must consider dependencies. There can be intra- and inter-temporal dependencies among the risk types. For example, in case of a major defect, the chance for a delay in the same period is higher than usual. We account for these dependencies between two risk types based on the Bravais-Pearson correlation coefficient ρ [35]. As an SDP lasts multiple periods or sprints, the risk types may also influence themselves over time. For example, if a major delay occurs, the chance for a greater delay in the following periods could be higher. Further, the risk of defect may tend to increase throughout the project in TSD mode. This drift that increases costs of defects over project time could be modeled via martingales. Alternatively, we could use individual distributions for every sprint, which would lead to more parameters and a more complex model. In our decision model, we therefore abstract from such auto-correlated structures, assuming that the distribution of each risk type is identical over all periods (Assumption 5). This assumption leads to an underestimation of the costs of defects in TSD mode. In general, assuming inter-temporal independence underestimates the overall risk exposure. However, as the risk is underestimated in both SD modes and as we compare both SD modes regarding the implementation of a concrete SDP, this has not effect on the recommendations of our decision model. On this assumption, we can sum up the expected additional costs and variances of the three risk types per sprint t to derive the total expected additional costs μ_R and their variance σ_R^2. The result is shown in Eq. (6) and (7).

$$\mu_R = \sum_{t=1}^{n} \left(\mu_{t,\mathrm{DEF}} + \mu_{t,\mathrm{CHG}} + \mu_{t,\mathrm{DEL}}\right) = n \cdot \left(\mu_{\mathrm{DEF}} + \mu_{\mathrm{CHG}} + \mu_{\mathrm{DEL}}\right) \tag{6}$$

$$\sigma_R^2 = \sum_{t=1}^{n} \sum_{i} \sum_{j} \sigma_{t,i}\, \sigma_{t,j} \rho_{i,j} = n \cdot \sum_{i} \sum_{j} \sigma_i\, \sigma_j \rho_{i,j} \tag{7}$$

with $i, j \in \{\mathrm{DEF}, \mathrm{CHG}, \mathrm{DEL}\}$

On this foundation, the objective function from Eq. (2) can be rewritten as shown in Eq. (8). Thereby, C_{Dev} represents the costs for a software developer.

$$\Delta\Phi = -C_{\mathrm{Dev}} \cdot a \left(size^E - d_n \cdot size^E + n \cdot \left(\frac{size}{n} \cdot (1 + c) \right)^E \right) - \Delta\mu_R - \frac{\alpha}{2} \cdot \Delta\sigma_R^2 \tag{8}$$

4 Evaluation

To validate our decision model, we implemented its design specification as a software prototype [36]. The prototype enabled conducting sample calculations and a sensitivity analysis. As for our calculations, we consider two fictional sample SDPs specified in Table 1. This setting does of course not account for all possible projects. We chose the parameters such that one project favors each mode. The input parameters for the base costs are identical for both projects, but the risk parameters differ. Depending on the specific cost/risk profile, we have a TSD-type project with a high exposure to defects and an ASD-type project with a high exposure to requirement changes and delays. One example for such a TSD-type project is the implementation of a new core system within a bank. An example for the ASD-type project is the development of a new payment application for smartphones. For the TSD-type project, the expected additional costs of defect and the related variance are quite large, while the other risks are rather small. For the ASD-type project, it is the vice versa. Additionally, the risks of the TSD- and the ASD-type project vary according to the risk profiles of both SD modes. We keep the size flexible to enable a comparison between costs and risks for both SD modes with respect to different project sizes.

Table 2 shows the correlation coefficients of the risk types for both modes, which we use for our calculations. For both modes, the risk of defect is positively correlated with the risk of delay. As TSD is not as adaptive and flexible as ASD, a defect in TSD has a larger impact on the risk of delay compared to ASD. For TSD, a change of requirements increases the risk of delay, as heavy planning does not allow for changes. ASD mode can cope much better with changed requirements. As a consequence, the risk of requirement changes and risk of delay are uncorrelated in ASD mode.

Table 1. Specification of a TSD-type and an ASD-type sample project

Base costs input parameters				Risks TSD-type project			Risks ASD-type project		
a	2.94	SF_5	6.24		TSD mode	ASD mode		TSD mode	ASD mode
B	0.91	EM_i	1						
G	3.67	C_{Dev}	5,000 MU	μ_{DEF}	1.5 %	4 %	μ_{DEF}	1 %	2 %
D	0.28	c	0.15	μ_{CHG}	1.25 %	0.5 %	μ_{CHG}	2.5 %	1.5 %
SF_1	3.72	e	0.05	μ_{DEL}	1.25 %	0.5 %	μ_{DEL}	2.5 %	1.5 %
SF_2	3.04	f	0.4	σ_{DEF}	5 %	30 %	σ_{DEF}	5 %	15 %
SF_3	4.24	Sprints	2 per mo.	σ_{CHG}	15 %	10 %	σ_{CHG}	22.5 %	10 %
SF_4	4.38	$Size$	$size$	σ_{DEL}	15 %	10 %	σ_{DEL}	22.5 %	10 %

Table 2. Sample correlations between risk types for TSD and ASD mode

TSD mode				ASD mode			
$\rho_{i,j}$	DEF	CHG	DEL	$\rho_{i,j}$	DEF	CHG	DEL
DEF	1	0	0.5	DEF	1	0	0.25
CHG	0	1	0.5	CHG	0	1	0
DEL	0.5	0.5	1	DEL	0.25	0	1

We first analyze the ASD-type project from Table 1 and choose a constant absolute risk aversion $\alpha = 0.0002$. This is reasonable according to Bamberg and Spremann [37]. Figure 2 (left chart) illustrates the results, where we show different deltas between TSD and ASD mode in terms of costs, risk, and risk-adjusted costs while varying the project size between 0 and 300 KSLOC. The dotted line depicts the delta in terms of costs, the dashed line depicts the delta in terms of risk, and the continuous line depicts the delta in terms of risk-adjusted costs.

As outlined, the cost advantage for ASD is only given for small projects. The initial effort required for accomplishing a small number of sprints in ASD mode is relatively small compared to the overhead of upfront planning in TSD mode. With larger project sizes, the initial planning effort in TSD mode increases less than the overhead for planning larger amounts of sprints in ASD mode. Thus, TSD has a cost advantage for larger project sizes. This finding in our sample calculation complies with extant knowledge. Researchers have argued that TSD mode is more favorable for large projects in general [38]. From a risk perspective, ASD is the preferable SD mode for all investigated project sizes between 0 and 300 KSLOC. The risk advantage for ASD always exceeds the cost advantage for TSD mode. Therefore, ASD is the appropriate mode for the sample project at hand. For decision-makers with a lower level of risk aversion (α), the delta between TSD and ASD in terms of risk-adjusted costs is depicted by the continuous grey line in Figure 2 (left chart). In this case, ASD is not favorable for all project sizes. In case of more risk-seeking decision-makers, the cost advantage for TSD exceed the risk advantage for ASD for large project sizes. This example illustrates that deciding on the SD mode is similar to balancing the certain cost advantage of TSD against the risk advantage of ASD. Economically speaking, decision-makers charge a risk premium for requirements changes on the safe cost advantage for TSD mode.

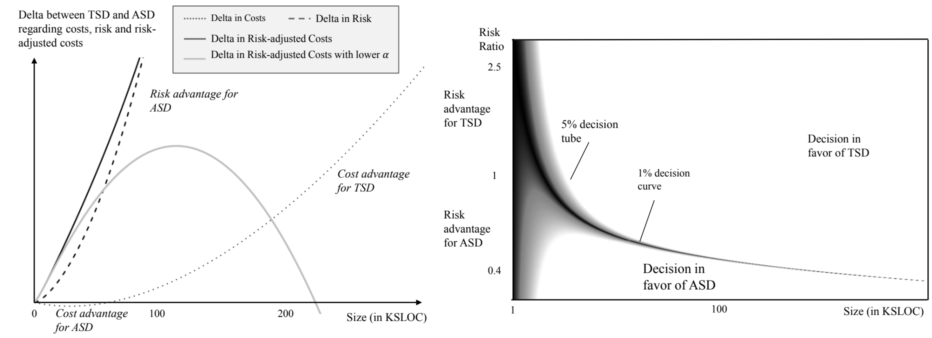


Figure 2. Calculation for the ASD-type sample project and grey-scale plot of relative delta

Next, we analyze the TSD-type project from Table 2. Table 1 shows the correlation coefficients of the risk types for both modes, which we use for our calculations. For both modes, the risk of defect is positively correlated with the risk of delay. As TSD is not as adaptive and flexible as ASD, a defect in TSD has a larger impact on the risk of delay compared to ASD. For TSD, a change of requirements increases the risk of delay, as heavy planning does not allow for changes. ASD mode can cope much better with changed requirements. As a consequence, the risk of requirements changes and risk of delay are uncorrelated in ASD mode. The cost advantages of TSD and ASD are the same as for ASD-type project, as the input parameters for the base costs are independent of the SD mode. Further, the risk advantage of TSD is positive for all project sizes, making TSD the preferable mode for nearly all project sizes. With the above examples, we demonstrated the suitability of our decision model by conducting sample calculations for a TSD- and an ASD-type sample project.

To demonstrate the robustness of our decision model, we now analyze the delta in risk-adjusted costs between both SD modes by varying the size and risk structure of a given project using a sensitivity analysis. When setting up the sensitivity analysis, two features must be considered: First, to stay comparable with an increasing project size, we rely on the relative delta in risk-adjusted costs, i.e., we divide the absolute delta in risk-adjusted costs by the absolute risk-adjusted costs of the respective SD mode. We do not consider the absolute risk-adjusted costs delta, which gets biased with an increasing project size. Second, as we only have a single dimension for describing the risk structure, we divide the risk of defect, which is typical for TSD-type projects, by the sum of the risk of requirement changes and delay, which are more typical for ASD-type projects. Thereby, we get a risk ratio that indicates a risk profile in favor of ASD for values from the interval $[0,1[$ and in favor of TSD for values from $]1,\infty[$. As a result, we get a three-dimensional analysis with project size and risk ratio on the horizontal and vertical axis, and the delta in relative risk-adjusted costs as a grey-scale plot. The delta in relative risk-adjusted costs is depicted on a grey-scale with bright indicating a high delta and dark indicating a low delta. The dark area shows settings where decision-makers are indifferent, whereas the bright areas illustrate robust decisions in favor of either TSD or ASD. We start with the TSD-type project from Table 1. We then reduce the risk of defect successively, while equivalently increasing the risk of requirement changes and delay until we end up with the ASD-type project from Table 1. This way, we get comparable projects regarding costs and a balanced risk profile half way between both projects (i.e., risk ratio = 1). The black area and line in Figure 2 (right chart) shows the decision curve, i.e., the area of indifference between both SD modes. Along the decision curve, the relative delta in the risk-adjusted costs is less than one percent. The grey decision tube around the black line depicts the delta values that are less than five percent of the respective risk-adjusted costs. Within this tube, the decision is not profound, but slightly tends to one of both modes. Decision-makers have to be careful regarding decisions within the five percent decision tube, as estimation errors may strongly influence the decision. For small projects, the decision is mainly influenced by the risk advantage of TSD or ASD depending on the SD mode, since the cost advantage of ASD is negligibly small. The black area indicates that our decision model is indifferent for small projects as it strongly depends on the risk distribution of the project at hand and the SD modes. However, this indifference is plausible, as the decision on the appropriate SD mode for a small SDP does not carry as much weight in terms of total costs as for large SDPs. The managerial implication is that decision makers should focus on finding the appropriate mode for larger SDPs, since for small projects the cost and risk difference is negligibly small. With an increasing project size, the costs are strongly in favor of TSD mode, leading to a steep decrease of the decision curve. The decision curve then converges towards the horizontal axis, indicating a larger area for unambiguous decisions in favor of TSD for very large projects. This is due to the cost advantage of TSD. That is, it can be reasonable from a cost/risk perspective to implement an SDP in TSD despite a risk advantage for ASD mode. For strongly ASD-type SDPs, i.e., projects with a very strong risk of requirement changes and delays, however, ASD is appropriate. The managerial implication of this finding is that decision makers should carefully evaluate the project's risk in both modes, since it could reverse a strong cost advantage for large projects. Although the graph in Figure 2 indicates otherwise, there is no reason to believe that the majority of projects should be conducted in TSD mode. First, SDPs in practice are not equally distributed over the size and risk dimensions. Second, we analyzed only one particular sample SDP in Figure 2. In a practical setting, our model primarily serves as decision support. Decision-makers should nevertheless carefully evaluate the outcome by taking their experience and other models into account.

5 Conclusion

In this study, we investigated how organizations can decide whether to implement a distinct SDP in TSD or ASD mode. Building on the characteristics of the SDP in focus and the characteristics of both SD modes, we proposed a decision model that analyzes the costs and risks associated with the implementation of a distinct SDP. Our model builds on the cost estimation method COCOMO II for TSD mode and the extensions proposed by Benediktsson et al. [24] for ASD mode to achieve comparability between both SD modes. To extend a purely cost-based view, the model accounts for three major risk types related to the implementation of an SDP, i.e., risk of defect, requirement changes, and delay. Besides these risk types, the decision model incorporates sprint length and overhead costs as characteristics of TSD and ASD mode. Our contribution is twofold. First, we bring the two cost estimation approaches together. Second, we extend the solely cost based view by a risk perspective. As for evaluation, we applied the decision model to sample projects with different input parameters. We also conducted a sensitivity analysis based on a software prototype to validate the decision model's suitability and robustness. The sensitivity analysis corroborated that the decision model yields plausible results. Overall, we contribute to the prescriptive knowledge on software development with a decision model as concrete artifact and its instantiation as a software prototype.

Our decision model is beset with limitations. First, we assume the risk types to be independent over time and normally distributed. Thus, the decision model underestimates risks associated with implementing an SDP. Risks may also increase or decrease over time, e.g., the expected additional costs for a change in requirements may increase over time if the project is executed in TSD mode. Research on inter-temporal dependencies of SD risks is required to implement these effects into our decision model. Future research should explore the risk structure of SDPs. Second, the decision model focuses on costs and risks that accrue during SDP implementation. Considering an SDP's business value also requires integrating benefits. In ASD mode, design and run time can no longer be separated as benefits accrue after each sprint. In TSD mode, however, benefits only realize after the SDP has been fully completed. Thus, we recommend that further research extends the decision model toward a cost-benefit analysis that includes benefits and a runtime perspective. Third, we focused on individual SDPs, ignoring the portfolio perspective. Future research should investigate how project dependencies influence the TSD/ASD trade-off. Fourth, as a first step, we demonstrated and analyzed the decision model via sample projects and a sensitivity analysis. Thus, the decision model would benefit from naturalistic evaluation, e.g., real-world case studies.

6 Acknowledgement

This research was (in part) carried out in the context of the Project Group Business and Information Systems Engineering of the Fraunhofer Institute of Applied Information Technology FIT.

References

1. Horlach, B., Drews, P., Schirmer, I.: Bimodal IT: Business-IT Alignment in the Age of Digital Transformation. In: Multikonferenz Wirtschaftsinformatik (MKWI) 2016. Universitaetsverlag Ilmenau, Ilmenau (2016)
2. Gartner: Gartner Executive Programs Survey of More Than 2,300 CIOs Reveals Many Are Unprepared for Digitalization: the Third Era of Enterprise IT. (2014)

3. Aitken, A., Ilango, V.: A Comparative Analysis of Traditional Software Engineering and Agile Software Development. In: Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS), pp. 4751-4760. IEEE, Wailea, Maui, Hawaii (2013)
4. Sondergaard, P.: Digital Business Demands Bimodal Capabilities. ITNOW 58, 8-9 (2016)
5. Ahimbisibwe, A., Cavana, R.Y., Daellenbach, U.: A contingency fit model of critical success factors for software development projects: A comparison of agile and traditional plan-based methodologies. Journal of Enterprise Information Management 28, 7-33 (2015)
6. Jorgensen, M., Shepperd, M.: A Systematic Review of Software Development Cost Estimation Studies. IEEE Transactions on Software Engineering 33, 33-53 (2007)
7. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly 37, 337-355 (2013)
8. Peffers, K., Tuunanen, T., Rothenberger, M., Chatterjee, S.: A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24, 45-77 (2007)
9. Sommerville, I.: Software Engineering. Pearson, New York (2015)
10. Lee, G., Xia, W.: Toward Agile: An Integrated Analysis of Quantitative and Qualitative Field Data on Software Development Agility. MIS Quarterly 34, 87-114 (2010)
11. Conboy, K.: Agility from first principles: reconstructing the concept of agility in information systems development. Information Systems Research 20, 329-354 (2009)
12. Nerur, S., Mahapatra, R., Mangalaraj, G.: Challenges of migrating to agile methodologies. Communications of the ACM 48, 72-78 (2005)
13. Dybå, T., Dingsøyr, T.: Empirical studies of agile software development: A systematic review. Information and Software Technology 50, 833-859 (2008)
14. Royce, W.W.: Managing the development of large software systems. In: Proceedings of IEEE WESCON. IEEE, Los Angeles (1970)
15. Beck, K., Beedle, M., Bennekum, A.v., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R.C., Mellor, S., Schwaber, K., Sutherland, J., Thomas, D., http://agilemanifesto.org/ (Accessed: 11/01/2016)
16. Fridgen, G., Klier, J., Beer, M., Wolf, T.: Improving Business Value Assurance in Large-Scale IT Projects: A Quantitative Method Based on Founded Requirements Assessment. ACM Transaction on Management Informatoin Systems 5, 1-17 (2014)
17. Remer, D.S., Nieto, A.P.: A compendium and comparison of 25 project evaluation techniques. Part 1: Net present value and rate of return methods. International Journal of Production Economics 42, 79-96 (1995)
18. Clermont, P.: Cost-Benefit Analysis: It's Back in Fashion, Now Let's Make It Work. Information Strategy: The Executive's Journal 18, 6-11 (2002)
19. Graham, D.A.: Cost-benefit analysis under uncertainty. The American Economic Review 71, 715-725 (1981)
20. Bannerman, P.L.: Risk and risk management in software projects: A reassessment. Journal of Systems and Software 81, 2118-2133 (2008)
21. Boehm, B.W., Clark, B., Horowitz, E., Westland, C., Madachy, R., Selby, R.: Cost models for future software life cycle processes: COCOMO 2.0. Annals of Software Engineering 1, 57-94 (1995)
22. Cao, L., Ramesh, B., Abdel-Hamid, T.: Modeling dynamics in agile software development. ACM Transactions on Management Information Systems 1, 1-26 (2010)
23. Yang, D., Wan, Y., Tang, Z., Wu, S., He, M., Li, M.: COCOMO-U: An Extension of COCOMO II for Cost Estimation with Uncertainty. In: Wang, Q., Pfahl, D., Raffo, D., Wernick, P. (eds.) Software Process Change, vol. 3966, pp. 132-141. Springer, Berlin, Heidelberg (2006)

24. Benediktsson, O., Dalcher, D., Reed, K., Woodman, M.: COCOMO-Based Effort Estimation for Iterative and Incremental Software Development. Software Quality Journal 11, 265-281 (2003)
25. Schmidt, R., Lyytinen, K., Mark Keil, P.C.: Identifying software project risks: An international Delphi study. Journal of Management Information Systems 17, 5-36 (2001)
26. Boehm, B.W.: Software risk management: principles and practices. IEEE Software 8, 32-41 (1991)
27. Fridgen, G., Mueller, H.-V.: Risk/cost valuation of fixed price IT outsourcing in a portfolio context. In: Proceedings of the 30th International Conference on Information Systems (ICIS), pp. 1115-1131. Association for Information Systems (AIS), Phoenix (2009)
28. Conboy, K., Coyle, S., Xiaofeng, W., Pikkarainen, M.: People over Process: Key Challenges in Agile Development. IEEE Software 28, 48-57 (2011)
29. Buhl, H.U., Röglinger, M., Stöckl, S., Braunwarth, K.S.: Value orientation in process management - Research Gap and Contribution to Economically Well-Founded Decisions in Process Management. Bus Inf Syst Eng 3, 163-172 (2011)
30. Zimmermann, S., Katzmarzik, A., Kundisch, D.: IT Sourcing Portfolio Management for IT Services Providers-A Risk/Cost Perspective. In: Proceedings of 29th International Conference on Information Systems (ICIS), pp. 1273-1290. Association for Information Systems (AIS), Paris (2008)
31. Pratt, J.W.: Risk aversion in the small and in the large. Econometrica 32, 122-136 (1964)
32. Shiohama, R., Washizaki, H., Kuboaki, S., Sakamoto, K., Fukazawa, Y.: Estimate of the appropriate iteration length in agile development by conducting simulation. In: Proceedings of the 2012 Agile Conference (AGILE), pp. 41-50. IEEE, Dallas (2012)
33. Brown, A.W., Ambler, S., Royce, W.: Agility at scale: economic governance, measured improvement, and disciplined delivery. Proceedings of the 2013 International Conference on Software Engineering, pp. 873-881. IEEE Press, San Francisco, CA, USA (2013)
34. Boehm, B.W., Turner, R.: Using risk to balance agile and plan-driven methods. Computer 36, 57-66 (2003)
35. Pearson, K.: Mathematical Contributions to the Theory of Evolution.--On a Form of Spurious Correlation Which May Arise When Indices Are Used in the Measurement of Organs. Proceedings of the Royal Society of London 60, 489-498 (1896)
36. Sonnenberg, C., Brocke, J.: Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research. In: Peffers, K., Rothenberger, M., Kuechler, B. (eds.) Proceedings of Design Science Research in Information Systems. Advances in Theory and Practice: 7th International Conference, DESRIST 2012, Las Vegas, NV, USA, pp. 381-397. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
37. Bamberg, G., Spremann, K.: Implications of constant risk aversion. Zeitschrift für Operations Research 25, 205-224 (1981)
38. Barlow, J.B., Keith, M.J., Wilson, D.W., Schuetzler, R.M., Lowry, P.B., Vance, A., Giboney, J.S.: Overview and Guidance on Agile Development in Large Organizations. Communications of the Association for Information Systems 29, 25-44 (2011)

Der Einfluss des Agilitätsgrads auf den Erfolg von Softwareentwicklungsprojekten unter Berücksichtigung der Unternehmenskultur

Martin Haase[1], Jan Jöhnk[2], Stefan Lipowsky[1] und Nils Urbach[2]

[1] it-economics GmbH, München, Deutschland
{mhaase, slipowsky}@it-economics.de
[2] Projektgruppe Wirtschaftsinformatik des Fraunhofer FIT, Universität Bayreuth, Bayreuth, Deutschland
{jan.joehnk, nils.urbach}@uni-bayreuth.de

Abstract. Agile Methoden sind seit einigen Jahren in der Softwareentwicklung etabliert und erfahren vor dem Hintergrund digitaler Transformationsprojekte aktuell eine sehr hohe Aufmerksamkeit. Gleichzeitig erscheinen der Zusammenhang zwischen dem gewählten Agilitätsgrad eines Softwareentwicklungsprojekts und seinem Erfolg sowie die moderierende Rolle der Unternehmenskultur weitestgehend ungeklärt. An dieser Stelle setzt unsere Forschung an, indem sie die Erfolgswirkung von Agilität in Softwareprojekten quantitativ-empirisch untersucht und dabei die spezifische Kultur des jeweiligen Unternehmens berücksichtigt. Hierzu wurden mit Hilfe eines Fragebogens insgesamt 108 verwertbare Antworten von IT-Projektmitarbeitern gesammelt und mit dem Ansatz Partial Least Squares in einer Mehrgruppenanalyse statistisch ausgewertet. Die Ergebnisse zeigen, dass einerseits ein höherer Agilitätsgrad zu einem höheren Projekterfolg führt, diese Wirkung andererseits aber von der Unternehmenskultur beeinflusst wird. Entsprechend erweitern unsere Forschungsergebnisse die Wissensbasis in diesem Bereich und liefern wertvolle Erkenntnisse für die Praxis der Softwareentwicklung.

Keywords: Agile Softwareentwicklung, Projekterfolg, Unternehmenskultur

1 Einleitung

Nach nunmehr 15 Jahren seit dem Verfassen des agilen Manifests sind agile Methoden im Bereich der Softwareentwicklung fest etabliert [1]. Vor dem Hintergrund von Digitalisierungsaktivitäten in zahlreichen Unternehmen erfahren diese Methoden derzeit eine hohe Popularität. Trotzdem werfen agile Methoden und deren zugrundeliegenden Werte und Prinzipien nach wie vor Fragestellungen auf, die einer Beantwortung bedürfen, um ein klares Bild dieser Vorgehensweise zu erhalten und eine entsprechende adäquate Anwendung zu ermöglichen. Dies konstatiert auch Hunt [2] als einer der Autoren des Manifests beim Rückblick auf die Entwicklungen in diesem Bereich. Seiner Beobachtung nach sei der Grundgedanke der agilen

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Haase, M.; Jöhnk, J.; Lipowsky, S.; Urbach, N. (2017): Der Einfluss des Agilitätsgrads auf den Erfolg von Softwareentwicklungsprojekten unter Berücksichtigung der Unternehmenskultur, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 454-468

Softwareentwicklung verloren gegangen und Agilität werde oftmals auf die reine Anwendung von entsprechenden Methoden und Techniken beschränkt. Wie verschiedene Beobachtungen zeigen, werden agile Methoden in Unternehmenskontexten nicht selten unreflektiert eingeführt, ohne deren Eignung im Vorfeld zu prüfen [3]. Nicht immer scheinen agile Methoden vorteilhafter gegenüber den klassischen, plangetriebenen Ansätzen zu sein, auch wenn die gegenwärtigen Diskussionen dies teilweise suggerieren. Neben spezifischen Projekteigenschaften stellt vor allem die vorherrschende Unternehmenskultur einen wichtigen Einflussfaktor dar, der grundlegend beachtet werden sollte [4, 5].

Eine weitere Beobachtung sind die dauerhaft niedrigen Erfolgsquoten von IT-Projekten. So wird durch zahlreiche Studien, trotz der weitläufigen Kritik aufgrund variierender Erfolgsdefinitionen sowie fehlender Kontext- und Projektinformationen, doch deutlich, dass der Erfolg von IT-Projekten im Allgemeinen als vergleichsweise niedrig angesehen werden kann [6]. So wurde beispielweise im Jahr 2014 im Rahmen des jährlichen *Chaos Report* der Standish Group, eine Erfolgsquote von nur 16,2% aller IT-Projekte festgestellt [7]. Über Erfolg oder Misserfolg von Projekten entscheidet unter anderem eine adäquate Auswahl und Anwendung der Projekt-Vorgehensweise. Deren Erfolg ist wiederum von der Kompatibilität zur jeweiligen Unternehmenskultur abhängig [8].

Entsprechend erscheint eine genauere Untersuchung des Zusammenhangs von der gewählten Projekt-Vorgehensweise und dem Erfolg eines IT-Projekts unter Berücksichtigung der gegebenen Unternehmenskultur als wertvoll für ein besseres Verständnis im IT-Projektmanagement. Mit spezifischen Fokus auf agile Praktiken und Softwareentwicklungsprojekte möchten wir mit diesem Forschungspapier der Beantwortung der folgenden Forschungsfrage nachgehen: *Welchen Einfluss hat der Agilitätsgrad der eingesetzten Vorgehensweise auf den Erfolg eines Softwareentwicklungsprojekts unter Berücksichtigung der Unternehmenskultur?*

Zur Beantwortung dieser Forschungsfrage haben wir ein quantitativ-empirisches Forschungsdesign gewählt. Auf Basis früherer Forschungsarbeiten leiten wir zunächst ein konzeptionelles Forschungsmodell ab, welches wir im Anschluss mittels gewonnener Daten aus einer Umfrage unter IT-Projektmitarbeitern analysieren. Als statistischen Ansatz zur Auswertung des empirischen Datenmaterials und zur Validierung des Forschungsmodells setzten wir den varianzbasierten Ansatz Partial Least Squares (PLS) ein.

2 Theoretische Entwicklung

2.1 Agile Softwareentwicklung

Mit dem Entstehen von verschiedenen Softwareentwicklungsmethoden in den 90er Jahren wurde der Grundstein für die agile Vorgehensweise in der Softwareentwicklung gelegt. So existieren aktuell gängige agile Methoden wie *Scrum* oder *Extreme Programming* (XP) bereits seit diesem Zeitraum [1, 9]. Die Entstehung resultierte aus der Überzeugung heraus, dass sich bisherige, plangetriebene Vorgehensweisen wie die Wasserfallmethode nicht mehr für die wachsende Dynamik im

Softwareentwicklungsbereich eignen [10]. Durch diese bewusste Abgrenzung wurde im Jahr 2001 das agile Manifest verfasst, um die Kernelemente der entstandenen Methoden aufzunehmen und damit die Basis der neuen agilen Vorgehensweise zu schaffen [11]. Seit diesem Zeitpunkt hielt die agile Softwareentwicklung Einzug in Praxis und Wissenschaft und gilt heute als eine der zentralen Vorgehensweisen [12].

Obwohl mittlerweile verschiedenste Definitionen der agilen Softwareentwicklung existieren, wird das agile Manifest nach wie vor am Häufigsten zur Charakterisierung herangezogen [13]. Diese Definition besteht dabei aus vier Werten und zwölf Prinzipien, worauf sich wiederum konkrete agile Methoden wie *Scrum* sowie entsprechend genutzte Praktiken begründen [11]. Die vier Werte stellen die Leitsätze der agilen Vorgehensweise dar. Dabei werden in erster Linie der Mensch und dessen Rolle im Softwareentwicklungsprozess in den Vordergrund gestellt. Besonders die enge Interaktion innerhalb des Entwicklungsteams sowie die Kollaboration mit dem Auftraggeber sind dabei essentiell. Außerdem wird die frühzeitige Auslieferung von Software-Funktionalitäten in möglichst kurzen Abständen angestrebt. Um dabei die kontinuierliche Aufnahme und Bearbeitung von neuen oder geänderten Anforderungen zu gewährleisten, wird eine schnelle Reaktionsfähigkeit während des gesamten Softwareentwicklungsprozesses vorausgesetzt. Dafür werden weiterhin eine kontinuierliche Informationsverfügbarkeit sowie ein kompetentes und eigenverantwortliches Team gefordert [14]. Die zwölf Prinzipien dienen wiederum als Operationalisierung, deren Berücksichtigung die Erfüllung der agilen Werte ermöglicht. Zusammenfassend wird dabei die Zufriedenstellung des Auftraggebers durch die Auslieferung qualitativer Software fokussiert. Darüber hinaus werden besonders Fähigkeiten der Selbstorganisation sowie der Anpassung bei Unsicherheiten aller Beteiligten hervorgehoben. Auch die kontinuierliche Verbesserung und Weiterentwicklung der Arbeitsweise während der Softwareentwicklung sowie direkte Kommunikationsflüsse werden angestrebt [15].

2.2 Projekterfolg

Der Projekterfolg wird seit langem immer wieder von der Forschung aufgegriffen, wobei die Anfänge bis in die 70er Jahre zurückreichen [16]. Dennoch konstatiert Prabhakar [17], dass weder in der Praxis noch in der Wissenschaft eine allgemeingültige Definition für den Erfolg von Projekten existiert. Diese Heterogenität schlägt sich auch auf die Erhebung des Projekterfolgs nieder, da je nach Erfolgsdefinitionen die Zielerreichung von IT-Projekten stark schwankt [6]. Trotz oder gerade aufgrund der verschiedenen beteiligten Stakeholder ist eine adäquate Abbildung des Projekterfolgs jedoch unerlässlich [18].

Traditionell wird der Projekterfolg anhand des *Iron Triangle*, dem Projektmanagementdreieck, dargestellt. Die Zielerreichung wird dabei in den Dimensionen Zeit, Kosten und Anforderungen gemessen [19]. Die Erreichung der geplanten Anforderungen ist im Vergleich zu den anderen beiden Dimensionen nicht klar abgrenzbar und wird häufig auch als Qualität, Leistung, Spezifikationen sowie Umfang bezeichnet [20]. Insbesondere im Hinblick auf die unterschiedlichen Perspektiven der beteiligten Stakeholder stellen diese drei Projekterfolgsdimensionen eine vereinfachte Abbildung des multidimensionalen Konstrukts dar und werden demnach oft als unzureichend kritisiert [u.a. 19]. Zudem wird mittels dieser Bestandteile nur die kurzfristige Sicht auf den Projekterfolg dargestellt, wobei längerfristige Auswirkungen durch das Projekt außen vor bleiben [21].

Daher existiert in der Literatur neben den dominierenden traditionellen Erfolgskriterien eine Vielzahl von weiteren Erfolgsbestandteilen und alternativen Projekterfolgsdefinitionen [20]. Dazu zählen die Zufriedenheit der beteiligten Stakeholder (Auftraggeber, Projektmitarbeiter und Nutzer), Steigerungen der Effektivität und Effizienz sowie der Beitrag zum Unternehmenserfolg [19, 20]. Dabei wird häufig zwischen Produkt- und Prozesserfolg unterschieden [18]. Der Prozesserfolg kann auch als Projektmanagementerfolg bezeichnet werden und umfasst die traditionellen Dimensionen des *Iron Triangle* [22]. Die Prozesseffizienz, welche das Verhältnis aus Zielerreichung und Aufwand beschreibt, wird jedoch teilweise als eigenständige Erfolgskomponente angesehen. Dies resultiert daraus, dass ein Projekt zwar hoch effizient durchgeführt werden kann und dennoch nicht die gesteckten Ziele erreicht, da diese bspw. von Anfang an unrealistisch waren [20]. Zum Produkterfolg wird dagegen die Erreichung der strategischen Unternehmensziele des Auftraggebers, des verfolgten Projektzwecks sowie die Bedürfnisse beteiligter Personengruppen gezählt [23]. Die Zufriedenheit der beteiligten Stakeholder kann entweder als Bestandteil oder aber auch losgelöst von den beiden Kategorien angesehen werden [23, 24]. Diese Überschneidungen der Kategorien sind neben verschiedenen Begriffsabgrenzungen insbesondere auf die Interdependenzen zwischen den einzelnen Projekterfolgsbestandteilen zurückzuführen. So ist bspw. die Erreichung von geplanten Anforderungen für das eigentliche Projektziel, die Erfüllung der Stakeholderbedürfnisse, zuständig. Dies bedingt wiederum die Zufriedenheit aller Beteiligten [25]. Auf Grundlage der diskutierten Literatur haben wir Projekterfolg im Rahmen dieser Arbeit als multidimensionales Konstrukt mit den Bestandteilen Prozesserfolg (Zielerreichung in Zeit, Kosten und Anforderungen; Prozesseffizienz) und Produkterfolg (Ergebnisqualität; Stakeholderzufriedenheit) definiert.

Seit den Anfängen der agilen Vorgehensweise werden die dazugehörigen Methoden zur Softwareentwicklung mit Erfolg assoziiert. Andere, plangetriebene Methoden seien demnach nicht mehr zeitgemäß und müssten innerhalb der Softwareentwicklung von Methoden anderer Vorgehensweisen abgelöst werden [26]. Wissenschaftliche Nachweise über die generelle Vorteilhaftigkeit der agilen Softwareentwicklung wurden bisher jedoch nur spärlich erbracht [27]. Trotzdem existieren verschiedene Veröffentlichungen, die Anhaltspunkte für die Vorteilhaftigkeit liefern. Dazu zählt die empirische Untersuchung von Serrador und Pinto [28]. Sie stellen fest, dass je agiler Projekte bewertet wurden, desto höher waren die Zufriedenheit der Beteiligten sowie die übergreifende Effizienz des Projekts. Weiterhin stellen Bermejo et al. [29] fest, dass die Einhaltung agiler Prinzipien einen positiven Einfluss auf den Erfolg der Softwareentwicklungsprojekte hat. Weiterhin liefern Praxisstudien Anhaltspunkte für einen positiven Einfluss der agilen Vorgehensweise auf den Erfolg von Softwareentwicklungsprojekten. So stellt der bereits erwähnte bekannte *Chaos Report* des Jahres 2014 dar, dass sowohl die kontinuierliche Auslieferung von Software als auch ein iteratives Vorgehen, zwei wesentliche Merkmale der agilen

Softwareentwicklung, sowie die verbundene Komplexitätsreduktion der Projekte die bisherige geringe Erfolgsquote anheben können [7]. Entsprechend formulieren wir unsere erste Hypothese:

H1: Ein höherer Agilitätsgrad führt zu höheren Erfolgsquoten von Softwareentwicklungsprojekten.

2.3 Unternehmenskultur

Im Kontext des Unternehmens wird die Kultur in der Fachliteratur vermehrt über kognitive Aspekte wie Überzeugungen, Werte und Annahmen definiert [30]. Schein [31] definiert Unternehmenskultur als Muster gemeinsamer Annahmen einer Gruppe, die Probleme äußerer Anpassung und innerer Integration so erfolgreich lösten, dass sie als gültig begriffen und an neue Mitglieder weitergegeben werden. Nach seinem Modell der drei Kulturebenen beinhaltet die unterste Ebene die als selbstverständlich angenommenen Grundannahmen, auf denen die beiden oberen Ebenen basieren. Diese oftmals unbewussten Annahmen umfassen bspw. Überzeugungen, Auffassungen, Gefühle und Gedanken. Darauf aufbauend folgt die Ebene der Normen und Werte, welche zum Teil bereits wahrnehmbar ist. Diese Ebene liefert den Mitgliedern einer Kultur Verhaltensrichtlinien für bestimmte Situationen, sowie Strategien, Ziele und Philosophien. Die oberste Ebene ist die der sichtbaren Artefakte und Symbole. Dazu zählen bspw. gemeinsame Verhaltensweisen, Sprache und Rituale sowie organisationale Prozesse und Strukturen.

Werte sind dabei für das im Folgenden beschriebene Untersuchungsmodell besonders relevant. Diese repräsentieren Präferenzen von gewissen Situationen gegenüber anderen, die implizit erlernt und als selbstverständlich wahrgenommen werden. Sie liefern Aufschluss über Verhaltensweisen, Prinzipien und Eigenschaften, die in einer Kultur als wertvoll angesehen werden [32]. Nach Schein [31] können die Mitglieder über Werte diskutieren und diese akzeptieren – womit sie zur Grundlage der Kultur werden – oder ablehnen. Durch Reduzierung der Kulturdefinition auf wenige Elemente wird ein direkter Vergleich zwischen Kulturtypen ermöglicht [33]. Demnach erfüllt die verwendete Kulturdefinition von Schein bzw. das zugrundeliegende Modell den eigentlichen Zweck der vorliegenden Arbeit.

Unternehmenskultur ist immer eine Koexistenz verschiedener (Sub-) Kulturen [34]. Dabei unterschieden sich diese bspw. je Hierarchiestufe, je Abteilung oder sogar je Team [35]. Weiterhin können sich die Subkulturen aufgrund von Aufgaben, Rollen oder spezifischen Machtstrukturen voneinander unterscheiden [36]. Die jeweiligen Mitglieder sind nicht an eine Subkultur gebunden, sondern können gleichzeitig mehreren angehören [33]. Trotz der unterschiedlichen Perspektiven und Werte existieren oft verbindende Charakteristika, die alle Subkulturen des Unternehmens gemeinsam haben. Dazu zählt bspw. die geteilte Ansicht über die Identität des Unternehmens [37]. Für den empirischen Teil dieser Untersuchung wird die jeweilige Kultur des entsprechenden Projektumfeldes untersucht.

Die Unternehmenskultur stellt im IT-Kontext einen wesentlichen Forschungsgegenstand dar [38]. Besonders in der Softwareentwicklung scheint eine geeignete Unternehmenskultur von hoher Relevanz zu sein. So stellt Ahimbisibwe [39] fest, dass die Unternehmenskultur ein kritischer Erfolgsfaktor von Softwareentwicklungsprojekten ist. Besonders im spezifischen Kontext der agilen Softwareentwicklung nimmt die Unternehmenskultur eine zentrale Rolle ein. Im Gegensatz zur Arbeit von Chow und Cao [40], in der kulturelle Aspekte als nicht erfolgskritisch für den Einsatz agiler Methoden identifiziert wurden, weisen viele Autoren auf die Relevanz einer geeigneten Kultur hin. So stellen Strode et al. [5] eine Verbindung zwischen der Nutzung agiler Methoden und kulturellen Merkmalen, wie bspw. Zusammenarbeit und Loyalität her. Auch Cohen [41] beschreibt die Agilität primär als ein kulturelles Thema. Demnach kann ein Unternehmen nicht agil sein, wenn die entsprechende Unternehmenskultur nicht vorliegt. Weiterhin führt Nerur [4] die Unternehmenskultur und verbundene Faktoren wie bspw. Führungsstil und Teamwork als Schlüsselelemente zur Einführung agiler Methoden auf. Weitere Autoren weisen auf notwendige Charakteristika einer agilen Unternehmenskultur hin. So kommen Siakas und Siakas [32] zu dem Schluss, dass die Werte Anpassungsfähigkeit, Kollaboration im Team und mit Kunden sowie eine möglichst schnelle Auslieferung von Software durch die Unternehmenskultur gefördert werden sollte. Da verschiedene Kulturtypen mit unterschiedlichen Merkmalsausprägungen existieren, sollten sich diese demnach auch unterschiedlich gut für den Einsatz agiler Softwareentwicklung eignen [42, 43]. Entsprechend formulieren wir unsere zweite Hypothese:

H2: Die Unternehmenskultur hat einen moderierenden Einfluss auf den Zusammenhang zwischen Agilitätsgrad und Erfolg von Softwareentwicklungsprojekten.

3 Untersuchungsmethodik

Für die empirische Untersuchung der aufgestellten Hypothesen wurde ein quantitativ-empirischer Forschungsansatz gewählt. Konkret haben wir ein Fragebogen-basiertes Vorgehen zur Validierung angewendet. Die gewählte Analyseeinheit des Forschungsvorhabens ist dabei stets ein konkretes Projekt. Der Fragebogen richtet sich an IT-Projektmitarbeiter mit verschiedenen Rollen und Aufgabenbereichen.

3.1 Konstruktoperationalisierung

Für die Konstruktoperationalisierung haben wir – soweit möglich – auf etablierte Messmodelle zurückgegriffen. Aufgrund der vergleichsweise hohen Neuartigkeit der verwendeten Konstrukte mussten jedoch einige Operationalisierungen neu entwickelt werden. Eine Übersicht der verwendeten Messmodelle ist Tabelle 1 zu entnehmen.

Da im Rahmen dieser Forschungsarbeit das agile Manifest als Definition zugrunde gelegt wurde, ist vor allem die Erfüllung der zwölf Prinzipien für einen hohen Agilitätsgrad im Projekt essentiell [11]. Entsprechend haben wir den Agilitätsgrad als formatives Second-Order-Konstrukt mit zwölf Higher-Order-Konstrukten (HOC) modelliert. Jedes HOC, welches entsprechend ein agiles Prinzip repräsentiert, wurde anhand von jeweils zwei Indikatoren, welche die entsprechenden Lower-Order-Konstrukte (LOC) darstellen, gemessen [44]. Der jeweils erste Indikator spiegelt dabei das eigentliche Prinzip des agilen Manifests wieder, das zweite eine spezifische Konkretisierung des jeweiligen Prinzips auf Basis von Praktiken etablierter agiler Methoden.

Für die Projekterfolgsmessung konnten wir das Messinstrument an Forschungsarbeiten mit ähnlichen Forschungszielen anlehnen. Für die Erstellung des Messinstruments wurde daher zwischen Gesamterfolg, Prozesserfolg (Termintreue, Budgettreue, Zielerreichung, Prozesseffizienz) sowie Produkterfolg (Ergebnisqualität, Zufriedenheit des Projektteams, Kundenzufriedenheit) unterschieden. Dabei bestätigen die Arbeiten von Bermejo et al. [29] sowie Serrador und Pinto [28] die Anwendbarkeit der Erfolgsdimensionen Termintreue, Budgeteinhaltung, funktionale Zielerreichung und Prozesseffizienz. Die Indikatoren der Ergebnisqualität sowie der Zufriedenheit relevanter Stakeholder wurden Baccarini [23] entlehnt, wobei auch die aktuelle Forschung im agilen Umfeld diese Erfolgskriterien aufgreift [28, 29].

Zur Messung der vorherrschenden Unternehmenskultur haben wir auf das Competing Value Model (CVM) von Cameron und Quinn [35] aufgesetzt. Dieses theoretische Modell zählt zu den am häufigsten genutzten Modellen in der Unternehmenskulturforschung [45]. Zur Operationalisierung des CVM haben wir das darauf basierende Organizational Culture Assessment Instrument (OCAI) eingesetzt. Mit diesem Instrument lassen sich über sechs Dimensionen komplexe Kulturprofile des Unternehmens ableiten. Diese sechs Dimensionen umfassen dominante Charakteristika des Unternehmens sowie Führungsstile und Umgang mit Mitarbeitern. Weiterhin werden sowohl die strategische Ausrichtung und Erfolgskriterien als auch die vorhandene Bindungskraft im Unternehmen analysiert [35]. Jede der sechs Dimensionen wurde anhand von jeweils vier etablierten Indikatoren gemessen.

Tabelle 1. Konstruktoperationalisierung

Konstrukt	**Operationalisierung**	**Anzahl Indikatoren**	**Quelle(n)**
Agilitätsgrad	Zwölf Prinzipien des Agilen Manifests sowie jeweils eine Konkretisierung	24	[11, 46]
Projekterfolg	Gesamterfolg, Termintreue, Budgettreue, Zielerreichung, Prozesseffizienz, Ergebnisqualität, Zufriedenheit des Projektteams, Kundenzufriedenheit	8	[23, 28, 29]
Unternehmens-kultur	Arbeitsumfeld, Führung, Umgang mit Mitarbeitern, Bindungskraft, Strategische Ausrichtung, Erfolgskriterien	24	[35]

3.2 Datenerhebung

Zur Erhebung empirischer Daten haben wir eine Umfrage durchgeführt. Der eingesetzte Fragebogen umfasste die drei Hauptbereiche Agilitätsgrad, Projekterfolg und Unternehmenskultur. Darüber hinaus wurden weitere projekt- und personenbezogene Daten erhoben. Die Bewertung der Indikatoren zur Messung des

Agilitätsgrads und des Projekterfolgs erfolgte über fünfstufige Likert-Skalen. Zur Messung der Unternehmenskultur wurden ipsative Skalen genutzt. Dieser vergleichsweise untypische Skalentyp umfasst die Aufteilung einer festgelegten Punktzahl auf eine feste Anzahl an vordefinierten Antworten. Im Rahmen des OCAI wird dabei eine Aufteilung von 100 Punkten auf vier Antwortmöglichkeiten vorgegeben. Vor der eigentlichen Datenerhebung wurde der Fragebogen einer ausführlichen Prüfung unterzogen. Zum Pre-test des neuentwickelten Messinstruments für den Agilitätsgrad wurde ein Card-Sorting-Verfahren genutzt, welches u.a. von Kankanhalli et al. [47] vorgeschlagen wird. Dieses wurde mit drei Experten aus dem Softwareentwicklungsumfeld durchgeführt. Des Weiteren wurde ein Pilottest des Fragebogens mit vier weiteren Personen durchgeführt. In Einzelgesprächen wurden Anmerkungen erhoben und der Fragebogen entsprechend angepasst.

Die eigentliche Datenerhebung fand im Zeitraum vom Februar bis Mai 2016 statt. Auf verschiedenen Kanälen wurde die Umfrage unter Teilnehmern von Software-Entwicklungsprojekten beworben: auf den sozialen Netzwerk-Plattformen XING, LinkedIn und Twitter, im persönlichen Netzwerk bei Kunden und Mitarbeiter einer IT-Beratung sowie über den Newsletter einer Projektmanagementgesellschaft. Die Ansprache erfolgte vor allem über E-Mails sowie über die persönliche Kontaktaufnahme. Ziel war es, Projektbeteiligte unterschiedlicher Rollen (Projektleiter, Product Owner, Scrum Master, Entwickler etc.) zu erreichen und jeweils ein Projekt bewerten zu lassen. Eine Einschränkung auf eine spezifische Rolle wurde nicht vorgenommen. Die Begründung hierfür war, dass es je nach Unternehmen verschiedene Rollenbezeichnungen mit unterschiedlichen Aufgaben gibt und eine Abgrenzung die Stichprobe ungewollt in traditionelle bzw. agile Richtung verschoben hätte. In der vorliegenden Untersuchung wollten wir jedoch bewusst das gesamte Projektspektrum betrachten.

Die deskriptive Auswertung der Umfragedaten ergab, dass verschiedene Projektbeteiligte an der Befragung teilgenommen haben. Unter den 108 vollständigen Rückläufern sind die Antworten von 32 Projektmanagern, 31 Softwareentwicklern, eine geringere Anzahl agiler Rollen (13 Scrum Master und acht Product Owner) sowie 24 Teilnehmer andere Rollen. In Bezug auf die Projektcharakteristika ergab die Auswertung der Umfragedaten, dass Projekte aus der Bank- und Versicherungsbranche mit 70% klar dominieren. Weitere 10% der untersuchten Projekte stammen aus der Energieversorgungsbranche. Ebenfalls in der Stichprobe enthalten sind u.a. Projekte aus den Bereichen Forschung und Entwicklung, Informations- und Kommunikationstechnologie, Medien sowie Handel. Insgesamt konnten Mitarbeiter aus einer Vielzahl verschiedener Projekte gewonnen werden.

4 Analyse und Ergebnisse

Zur statistischen Analyse der gewonnenen Daten haben wir den varianzbasierten PLS-Ansatz und das Softwarepaket SmartPLS in der Version 3.0 genutzt. Unser konzeptionelles Modell wurde sowohl mit der gesamten Stichprobe als auch unterteilt nach dominierenden Kulturausprägungen analysiert. Da in der Stichprobe zwei

Kulturtypen nur vereinzelt dominierend vorlagen, wurden zwei Untersuchungsgruppen gebildet, die jeweils zwei Kulturtypen enthalten. In der Mehrgruppenanalyse repräsentiert Gruppe A den Stichprobenteil (n=55), in dem die Kulturtypen *Clan Culture* und *Adhocracy Culture* dominieren, während Gruppe B den Stichprobenanteil (n=53) umfasst, in dem *Hierarchy Culture* und *Market Culture* die dominierenden Kulturtypen darstellen. Bei dieser Einteilung wurden vor allem die Merkmale der einzelnen Kulturtypen beachtet. So beinhaltet Gruppe A Kulturtypen, die sich besonders durch Flexibilität und Dynamik auszeichnen, während die Kulturtypen in Gruppe B vorwiegend durch Stabilität und Kontinuität geprägt sind [42, 43].

4.1 Untersuchung der Messmodelle

Der erste Schritt der empirischen Analyse ist die Validierung der Messmodelle, welche wir sowohl formativ (Agilitätsgrad) als auch reflektiv (Projekterfolg) modelliert haben.

Für das zweidimensionale, formative Messmodell des Agilitätsgrads haben wir auf beiden Ebenen die Pfadkoeffizienten in Form der Pfadgewichte herangezogen, für die in der Fachliteratur Mindestwerte von 0,1 bzw. 0,2 gefordert werden [48]. Weiterhin wurde die Signifikanz dieser Pfadgewichte untersucht [49]. Auf der *LOC*-Ebene erreichten vorwiegend die Indikatoren, die direkt aus dem jeweiligen agilen Prinzip abgeleitet wurden, höhere Werte. In Verbindung mit der *HOC*-Ebene konnte weiterhin ein durchgängig signifikanter Einfluss auf den Agilitätsgrad ausgehend von sechs agilen Prinzipien ermittelt werden. Trotz teilweise fehlender Signifikanzen und schwacher oder negativer Pfadgewichte, haben wir das Messmodell – wie von mehreren Autoren empfohlen [u.a. 44] – aufgrund von inhaltlichen Überlegungen nicht nachträglich verändert. Besonders im Hinblick auf die zwölf agilen Prinzipien würde die Elimination einzelner Indikatoren bedeuten, dass eine anerkannte Definition nur teilweise durch das Untersuchungsmodell abgebildet werden würde. Für die Analyse wurde außerdem die Korrelation zwischen den Indikatoren mittels des Varianzinflationsfaktors (VIF) bewertet, um Multikollinearität zwischen den formativen Indikatoren auszuschließen. Dabei lagen alle Werte des Messmodells innerhalb des empfohlenen Wertebereiches von VIF<5 [44].

Das Konstrukt des Projekterfolges haben wir reflektiv modelliert. Zur Validierung wurden die Faktorladungen der einzelnen Indikatoren herangezogen, welche einen Mindestwert von $\lambda_i>0,7$ annehmen sollten [49]. Im Grundmodell erreichten sieben von acht Faktorladungen diesen Mindestwert. Lediglich ein Indikator lag mit $\lambda=0,697$ nur minimal – und damit vertretbar – darunter. Auch die in der Literatur geforderten Werte der Gütekriterien der Konstruktreliabilität ($0,6<\rho_c<1$), welche die interne Konsistenz durch Korrelation der Indikatoren bewertet, sowie die Konvergenzvalidität (0,5<AVE<1), die den erklärbaren Varianzanteil bemisst, wurden erfüllt [44].

Die Analysen der beiden kulturspezifischen Gruppen weisen Abweichungen vor allem innerhalb des Messmodells des Agilitätsgrads auf. Dabei wurden erforderliche Mindestwerte der Pfadgewichte sowie der verbundenen Signifikanzniveaus teilweise nicht erreicht. Dies ist u.a. durch die starke Verkleinerung der Stichprobe bei gleichbleibender Indikatoranzahl zu begründen, sollte bei der Interpretation der Ergebnisse aber dennoch berücksichtigt werden.

4.2 Untersuchung des Strukturmodells

Der zweite Schritt der empirischen Analyse stellt die Validierung des Strukturmodells dar. Ein wichtiges Gütekriterium sind dabei ebenfalls die Pfadkoeffizienten mit den oben beschriebenen Schwellenwerten [48]. Bei der Analyse der Gesamtstichprobe konnten sowohl die Höhe als auch die Signifikanz des Pfadkoeffizienten bestätigt werden. Ein weiteres Gütekriterium ist das Bestimmtheitsmaß (R^2), welches den Varianzanteil der endogenen Variablen, der durch die exogene Variable verursacht wird, wiedergibt. Der R^2-Wert von 0,54 für die abhängige Variable Projekterfolg ist als mittelhoch anzusehen [50]. Weiterhin wurde die Prognoserelevanz (Q^2) des Strukturmodells herangezogen, welche angibt inwieweit die empirischen Daten durch das Untersuchungsmodell besser rekonstruiert werden können, als über den Mittelwert. Der erreichte Q^2-Wert von 0,28 erfüllt den geforderten Mindestwert von $Q^2>0$ [50].

Die beiden kulturspezifischen Gruppen weisen ähnlich gute Werte für die Höhe und Signifikanz der Pfadkoeffizienten auf. Der signifikante Einfluss des Agilitätsgrads auf den Projekterfolg ist dabei für Gruppe A mit einem Wert von 0,86 am stärksten. Die Unterschiede zwischen den einzelnen Strukturmodellen werden aber vor allem bei der Betrachtung des Bestimmtheitsmaßes deutlich. Hierbei ist für Gruppe A ein Wert von $R^2=0{,}73$ feststellbar und ist damit als substanziell anzusehen. Bei Gruppe B verhält sich dieses Gütekriterium mit einem Wert $R^2=0{,}57$ ähnlich zur Gesamtstichprobe. Die weiterhin betrachtete Prognoserelevanz wird auch in den beiden kulturspezifischen Gruppen erfüllt, weist allerdings keine nennenswerten Unterschiede auf. Abbildung 1 fasst die Ergebnisse der Analyse auf Basis der Gesamtstichprobe sowie der beiden Untergruppen zusammen.

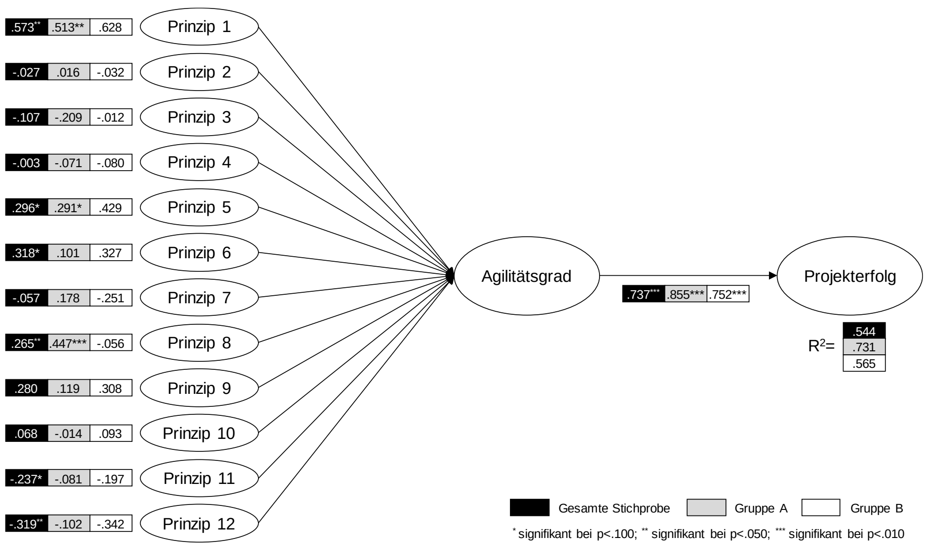


Abbildung 1. Empirische Validierung des Forschungsmodells

5 Diskussion der Ergebnisse

Die Ergebnisse der empirischen Validierung zeigen einen positiven Zusammenhang zwischen dem Agilitätsgrad und dem Projekterfolg, sowohl für die Gesamtichprobe als auch für die unterschiedlichen Kulturtypen. Dabei ist deutlich erkennbar, dass der Einfluss des Agilitätsgrads auf den Projeterfolg in der Gruppe A höher ist als in der Gruppe B. Dieses Ergebnis ist dahingehend zu interpretieren, dass bei einem höheren Agilitätsgrad eines Projektes, das in *Clan*- bzw. *Adhocracy Culture* durchgeführt wurde, auch der Projekterfolg höher ist. Das gleiche gilt zwar auch für Projekte in Unternehmen, in denen *Hierarchy Culture* und *Market Culture* die dominierenden Kulturtypen darstellen. Hier ist jedoch der Zusammenhang weniger stark ausgeprägt. Entsprechend unterstützt die empirische Untersuchung sowohl den hypothetisierten positiven Zusammenhang zwischen Agilitätsgrad und Projekterfolg von Softwareentwicklungsprojekten sowie den Einfluss der vorherrschenden Unternehmenskultur. Damit können beide zuvor aufgestellten Hypothesen bestätigt werden.

Aus den Ergebnissen ergeben sich verschiedene Implikationen für Wissenschaft und Praxis. Grundsätzlich kann auf Basis der analysierten Daten die Vorteilhaftigkeit einer agilen Vorgehensweise in Softwareentwicklungsprojekten festgestellt werden. Selbst über verschiedene Unternehmenskulturtypen hinweg waren positive Effekte auf den Projekterfolg erkennbar. Zusätzlich haben wir festgestellt, dass sich diese positiven Effekte zwischen den verschiedenen betrachteten Unternehmenskulturen hinsichtlich ihrer Stärke unterscheiden. Dieser Zusammenhang kann in Kombination mit weiterführenden Überlegungen und Analysen zur Bewertung der Vorteilhaftigkeit einer agilen Softwareentwicklung für den spezifischen Unternehmenskontext herangezogen werden. Eine Entscheidung für die agile Softwareentwicklung sollte allerdings von verschiedenen Faktoren abhängig gemacht werden. So wurde bereits im Grundlagenteil auf die Komplexität der Unternehmenskultur aufmerksam gemacht. Dabei bildet das genutzte *OCAI* nur einen Teil Unternehmenskultur ab. So können möglicherweise Faktoren, die nicht in dieser Analyse betrachtet wurden, einen negativen Einfluss auf die Nutzung der agilen Softwareentwicklung haben. Darüber hinaus ist auch Wandel hin zu einem anderen Kulturtyp ein langwieriger Prozess, der nicht unterschätzt werden sollte [35]. Demnach sollte grundsätzlich vorerst die jeweilige Unternehmenssituation betrachtet werden. Hierbei können bspw. weiterführende Kulturtests eingesetzt werden [51]. Der Erfolg agiler Softwareentwicklung ist nicht nur von der Unternehmenskultur abhängig. Demnach werden weitere erfolgskritische Rahmenbedingungen für eine erfolgreiche Einführung der agilen Softwareentwicklung in Betracht gezogen [4]. Ein Indiz dafür liefert auch die deskriptive Analyse unserer Daten. Hierbei wurden sehr viele Mischformen zwischen agilen und traditionellen Softwareentwicklungsmethoden festgestellt, was möglicherweise auch auf Barrieren für eine vollständige Agilitätsausprägung innerhalb der Unternehmen zurückzuführen ist. Weiterhin können auch die agilen Prinzipien selber ein Hindernis darstellen. So ist bspw. denkbar, dass ein iteratives Vorgehen weitreichende prozessuale Anpassungen nach sich zieht. Demnach sollte auch hier vorerst die Unternehmenssituation auf Eignung analysiert werden.

Trotz des nachgewiesenen Wirkungszusammenhangs sind die Limitationen unseres Forschungsvorhabens zu berücksichtigen. Im Rahmen der quantitativen Erhebung können verschiedene Verzerrungen aufgetreten sein, die es zu berücksichtigen gilt. Besonders der Social Desirablitiy Bias, der Verzerrungen aufgrund von sozialer Erwünschtheit umfasst, könnte die Ergebnisse beeinflusst haben [52]. So wird die agile Softwareentwicklung in der Fachliteratur nicht selten als generell vorteilhaft gegenüber der traditionellen Softwareentwicklung dargestellt [26]. Dadurch ist es grundsätzlich denkbar, dass die Umfrageteilnehmer dazu tendierten, Aussagen, die für die agile Softwareentwicklung sprachen, besser zu bewerten. Auch die Erhebung der Unternehmenskultur kann grundsätzlich durch diese Verzerrung betroffen sein. So ist eine hierarchische Kultur möglicherweise negativ behaftet [35]. Dies konnte allerdings nicht durch die Datenanalyse bestätigt werden, da dieser Kulturtyp am häufigsten dominierte. Weiterhin denkbar ist, dass auch Aussagen hinsichtlich des Projekterfolges verzerrt bewertet wurden [25]. Grundsätzlich konnten zwar auch sehr geringe Werte für den Projekterfolg ermittelt werden, diese bildeten allerdings die Ausnahme. Gleichzeitig war auch erkennbar, dass sehr hoch bewertete Projekterfolge ausblieben. Auch in Verbindung mit den Kulturtypen könnte eine Verzerrung aufgetreten sein, Demnach ist es denkbar, dass in Gruppe A agile Projekte als erfolgreicher bewertet wurden. Im Rahmen der Limitationsbetrachtung sollte auch die Datenanalyse berücksichtigt werden. Hierbei sind besonders die Bestandteile des Messmodells des Agilitätsgrads zu beachten. So konnten im Rahmen der Analyse die geforderten Werte für Signifikanz und Pfadgewichte teilweise nicht erfüllt werden. Eine Veränderung des Messmodells des formativen Konstrukts haben wir aus inhaltlichen Erwägungen dennoch abgelehnt. Grundsätzlich ist es deshalb möglich, dass auch die Zusammenhänge im Strukturmodell von diesen Effekten beeinflusst worden sind. Weiterhin ist anzumerken, dass für die kulturspezifische Betrachtung nur vergleichsweise kleine Stichproben genutzt werden konnten, was die Ergebnisse ebenfalls beeinflusst haben könnte.

6 Zusammenfassung und Fazit

Ziel der vorliegenden Forschungsarbeit war die empirische Untersuchung des Wirkungszusammenhangs vom Agilitätsgrad eines Softwareprojekts auf den Projekterfolg. Weiterhin sollte der moderierende Einfluss der vorherrschenden Unternehmenskultur berücksichtigt werden. Dazu haben wir ein deduktives Forschungsmodell durch eine Umfrage unter IT-Projektmitarbeitern validiert. Als statistischen Ansatz zur Auswertung des empirischen Datenmaterials haben wir die Strukturgleichungsmodellierungsmethode Partial Least Squares eingesetzt.

Die zu Beginn dieses Artikels dargestellte Forschungsfrage konnte durch den gewählten Forschungsansatz beantwortet werden. So zeigt sich, dass eine agile Softwareentwicklung besonders in einer geeigneten Unternehmenskultur erfolgreich ist. Aber auch unabhängig davon ist ein positiver Zusammenhang zwischen Agilitätsgrad und Projekterfolg feststellbar. Trotz der gewonnenen Erkenntnisse ist das Forschungsfeld nicht erschöpft. Vielmehr ergeben sich verschiedene

Anknüpfungspunkte für zukünftige Untersuchungen. Dabei weisen grundsätzlich alle drei Konstrukte verschiedene anerkannte Definitionen und Messmöglichkeiten auf. Besonders die Konzeption der Agilität sollte zukünftig wiederholt aufgegriffen werden, da das aktuell verwendete Messmodell Schwächen aufweist. So könnte bspw. die ebenfalls weit verbreitete Definition von Conboy [13, 53] zugrunde gelegt werden. Darüber hinaus stellt auch die kontextspezifische Interaktion zwischen Unternehmenskultur und agiler Vorgehensweise ein zukünftiges Forschungsfeld dar, um die Auswirkungen auf den Projekterfolg detailliert zu begründen [54]. Ein Vergleich der bisherigen mit zukünftigen Forschungsergebnissen kann dabei dazu beitragen, die Zusammenhänge der drei Konstrukte besser zu verstehen. Weiterhin sollte in weiteren Forschungsvorhaben die Stichprobengröße in Bezug auf die einzelnen Kulturtypen ausgeweitet werden. Dabei kann eine weitere Datensammlung und -analyse weiterhin zur Untersuchung von potenziellen Verzerrungsmöglichkeiten genutzt werden.

Literaturverzeichnis

1. Komus, A.: GPM-Studie Status Quo Agile 2014. Agiles und klassisches Projektmanagement Miteinander statt gegeneinander (2014)
2. Jackson, M.B.: Agile. A Decade in. PM Network 26, 58–62 (2012)
3. Grossman, F., Bergin, J., Leip, D., Merrit, S., Gotel, O.: One XP Experience: Introducing Agile (XP) Software Development into a Culture that is Willing but not Ready. In: Conference of the Centre for Advanced Studies on Collaborative research Proceedings, 242–254 (2004)
4. Nerur, S., Mahapatra, R.K., Mangalaraj, G.: Challenges of migrating to agile methodologies. Organizations must carefully assess their readiness before treading the path of agility. Communications of the ACM 48, 72–78 (2005)
5. Strode, D.E., Huff, S.L., Tretiakov, A.: The Impact of Organizational Culture on Agile Method Use. In: Hawaii International Conference on System Sciences Proceedings (2009)
6. Basten, D., Pankratz, O., Joosten, D.: Assessing the assessors. An overview and evaluation of it project success reports. In: ECIS 2013 Proceedings (2013)
7. The Standish Group: The Standish Group Report Chaos (2014)
8. Joslin, R., Müller, R.: Relationships between a project management methodology and project success in different project governance contexts. International Journal of Project Management 33, 1377–1392 (2015)
9. Abrahamsson, P., Warsta, J., Siponen, M.T., Ronkainen, J.: New Directions on Agile Methods. A Comparative Analysis. In: International Conference on Software Engineering Proceedings, 244–254 (2003)
10. Gren, L., Torkar, R., Feldt, R.: The prospects of a quantitative measurement of agility. Journal of Systems and Software 107, 38–49 (2015)
11. Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A. and Jeffries, R., et al.: Manifest für Agile Softwareentwicklung, http://www.agilemanifesto.org/iso/de/
12. Dingsøyr, T., Nerur, S., Balijepally, V., Moe, N.B.: A decade of agile methodologies. Towards explaining agile software development. Journal of Systems and Software 85, 1213–1221 (2012)

13. Hummel, M.: State of the Art. A Systematic Literature Review on Agile Information Systems Development. In: Hawaii International Conference on System Sciences Proceedings, 4712–4721 (2014)
14. Abrahamsson, P., Salo, O., Ronkainen, J., Warsta, J.: Agile software development methods. Review and analysis. VTT, Espoo (2002)
15. Trepper, T.: Fundierung der Konstruktion agiler Methoden. Anpassung, Instanziierung und Evaluation der Methode PiK-AS. Springer Fachmedien Wiesbaden, Wiesbaden (2015)
16. Ika, L.A.: Project success as a topic in project management journals. Project Management Journal 40, 6–19 (2009)
17. Prabhakar, G.P.: What is Project Success. A Literature Review. International Journal of Business and Management 3, 3–10 (2008)
18. Joosten, D., Basten, D., Mellis, W.: Measurement of Information System Project Success in German Organizations. International Journal of Information Technology Project Management 5, 1–20 (2014)
19. Shenhar, A.J., Levy, O., Dvir, D.: Mapping the dimensions of project success. The Professional Journal of the Project Management Institute 28, 5–13 (1997)
20. Pankratz, O. and Basten, D.: Ladder to Success – Eliciting Project Managers' Perceptions of IS Project Success Criteria (2014)
21. Atkinson, R.: Project management. Cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. International Journal of Project Management 17, 337–342 (1999)
22. Pankratz, O., Basten, D.: Do project characteristics influence the relevance of is project success dimensions? In: ECIS 2012 Proceedings (2012)
23. Baccarini, D.: The Logical Framework Method for Defining Project Success. Project Management Journal 30, 25–32 (1999)
24. Nelson, R.R.: Project Retrospectives. Evaluating Project Success, Failure, and Everything in Between. MIS Quarterly Executive 4, 361–372 (2005)
25. Thomas, G., Fernández, W.: Success in IT projects. A matter of definition? International Journal of Project Management 26, 733–742 (2008)
26. Highsmith, J., Cockburn, A.: Agile software development. The business of innovation. Computer 34, 120–127 (2001)
27. Lee, G., Xia, W.: Toward agile. An integrated analysis of qantitative and qualitative field data on software development agility. MIS Quarterly 34, 87–114 (2010)
28. Serrador, P., Pinto, J.K.: Does Agile work? A quantitative analysis of agile project success. International Journal of Project Management 33, 1040–1051 (2015)
29. Bermejo, P.H., Zambalde, A.L., Tonelli, A.O., Souza, S.A., Zuppo, L.A., Rosa, P.L.: Agile Principles and Achievement of Success in Software Development. A Quantitative Study in Brazilian Organizations. Procedia Technology 16, 718–727 (2014)
30. Sackmann, S.A.: Culture and Subcultures. An Analysis of Organizational Knowledge. Administrative Science Quarterly 37, 140–161 (1992)
31. Schein, E.H.: Organizational Culture and Leadership. John Wiley & Sons, San Francisco, USA (2010)
32. Siakas, K.V., Siakas, E.: The agile professional culture. A source of agile quality. Software Process: Improvement and Practice 12, 597–610 (2007)
33. Richter, T.: Kulturorientierte Forschung in der Wirtschaftsinformatik. Entwicklung eines Werkzeugs zur Abgrenzung kultureller Forschungskontexte und zur Ermittlung kontextuell passender Kulturbeschreibungsmodelle. lulu, Bonn (2014)
34. Tolfo, C., Wazlawick, R.S.: The influence of organizational culture on the adoption of extreme programming. Journal of Systems and Software 81, 1955–1967 (2008)

35. Cameron, K.S., Quinn, R.E.: Diagnosing and Changing Organizational Culture. Based on the Competing Values Framework. Wiley, Somerset, England (2006)
36. Scholz, C.: Personalmanagement. Informationsorientierte und verhaltenstheoretische Grundlagen. Vahlen, München (2000)
37. Whetten, D.A.: Albert and Whetten Revisited. Strengthening the Concept of Organizational Identity. Journal of Management Inquiry 15, 219–234 (2006)
38. Leidner, D.E., Kayworth, T.: A Review of Culture in Information Systems Research. MIS Quarterly 30, 357–399 (2006)
39. Ahimbisibwe, A., Cavana, R.Y., Daellenbach, U.: A contingency fit model of critical success factors for software development projects. Journal of Enterprise Information Management 28, 7–33 (2015)
40. Chow, T., Cao, D.-B.: A survey study of critical success factors in agile software projects. Journal of Systems and Software 81, 961–971 (2008)
41. Cohen, D., Lindvall, M., Costa, P.: An Introduction to Agile Methods. In: Zelkowitz, M.V. (ed.) Advances in Computers. Advances in Software Engineering, S. 1–66. Academic Press, Orlando (2004)
42. Maximini, D.: The scrum culture. Introducing agile methods in organizations. Springer, Cham (2015)
43. Iivari, J., Huisman, M.: The relationship between organizational culture and the development of systems developments methodologies. MIS Quarterly 31, 35–58 (2007)
44. Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M.: A primer on partial least squares structural equation modeling (PLS-SEM). Sage, Los Angeles (2016)
45. Yu, T., Wu, N.: A Review of Study on the Competing Value Framework. International Journal of Business and Management 4 (2009)
46. Williams, L.: Agile Software Development Methodologies and Practices. In: Zelkowitz, M.V. (ed.) Advances in computers, 80, S. 1–44. Academic Press, London (2010)
47. Kankanhalli, A., Tan, B.C.Y., Wei, K.-K.: Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation. MIS Quarterly 29, 113–143 (2005)
48. Ringle, C.M., Spreen, F.: Beurteilung der Ergebnisse von PLS-Pfadanalysen. Wirtschaftswissenschaftliches Studium 36, 211–216 (2007)
49. Nitzl, C.: Eine anwendungsorientierte Einführung in die Partial Least Square (PLS)-Methode. In: Hansmann, K.-W. (ed.) Industrielles Management. Arbeitspapier, S. 1–72. Universität Hamburg, Hamburg (2010)
50. Chin, W.W.: The Partial Least Squares Approach to Structural Equation Modeling. In: Marcoulides, G.A. (ed.) Modern Methods for Business Research, S. 295–358. Lawrence Erlbaum Associates, Mahwah, New Jersey (1998)
51. Jung, T., Scott, T., Davies, H.T.O., Bower, P., Whalley, D., McNally, R., Mannion, R.: Instruments for Exploring Organizational Culture. A Review of the Literature. Public Administration Review 69, 1087–1096 (2009)
52. Bhattacherjee, A.: Social science research. Principles, Methods, and Practices. USF Tampa Library Open Access Collections, Tampa, Florida (2012)
53. Conboy, K.: Agility from First Principles. Reconstructing the Concept of Agility in Information Systems Development. Information Systems Research 20, 329–354 (2009)
54. Schmidt, C.T., Kude, T., Heinzl, A., Mithas, S.: How Agile Practices Influence the Performance of Software Development Teams: The Role of Shared Mental Models and Backup. In: ICIS 2014 Proceedings, 1–18 (2014)

Bewertung und Planung von IT-Investitionen unter Berücksichtigung finanzieller Beschränkungen

Stefan Pfosser

Universität Augsburg,
Kernkompetenzzentrum Finanz- & Informationsmanagement,
86159 Augsburg, Deutschland
stefan.pfosser@fim-rc.de

Abstract. Große Investitionen in Informationstechnologie (IT) sind oftmals langfristig geplante Investitionsvorhaben, die sich über einen längeren Zeitraum erstrecken und dadurch Investitionsauszahlungen in mehreren Perioden erfordern. Bei deren Planung und Budgetierung muss daher ein mehrperiodiger Planungshorizont inklusive relevanter finanzieller Rahmenbedingungen und Risiken berücksichtigt werden. Dadurch können Höhe und Zeitpunkt der Investitionsauszahlungen über mehrere Perioden hinweg optimal bestimmt und der erhoffte Wertbeitrag des Investitionsvorhabens abgesichert bzw. ggf. sogar gesteigert werden. Dies gilt insbesondere für Unternehmen, deren verfügbare finanzielle Eigen- und Fremdmittel stark begrenzt bzw. besonders unsicher sind. Deshalb sollten bei der Planung von langfristigen IT-Investitionen sowohl aktuelle, als auch zukünftig drohende finanzielle Beschränkungen berücksichtigt werden. Dies erfordert einen Disziplinen-übergreifenden Ansatz, der Erkenntnisse aus Finanzmanagement und IT-Management verbindet. In diesem Beitrag wird daher ein Optimierungsmodell entwickelt, bei dem relevante finanziellen Rahmenbedingungen bei der Planung eines IT-Investitionsvorhabens berücksichtigt werden, wodurch dessen Wertbeitrag gesteigert werden kann.

Keywords: IT-Investition; IT-Management; Investitionsbewertung; Investitionsstrategie

1 Einleitung und Motivation

IT-Investitionen sind durch die zunehmende Durchdringung von Unternehmen mit IT und die rapide fortschreitende Digitalisierung ein entscheidender Wettbewerbsfaktor für Unternehmen. Sie leisten einen wichtigen Beitrag zum langfristigen wirtschaftlichen Erfolg eines Unternehmens. Gleichzeitig können Unternehmen allerdings gezwungen sein, aufgrund finanzieller Beschränkungen geplante Budgets für langfristige IT-Investitionen zu reduzieren (vgl. bspw. [15], [26], [35]). Dadurch kann ein Unternehmen die Rückflüsse aus dem jeweiligen Investitionsvorhaben nicht bzw. nicht in geplanter Höhe realisieren, wodurch der dazugehörige Wertbeitrag erheblich reduziert wird oder sogar ein negativer Wertbeitrag resultieren kann. Die finanzielle

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Pfosser, S. (2017): Bewertung und Planung von IT-Investitionen unter Berücksichtigung finanzieller Beschränkungen, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 469-483

Situation eines Unternehmens spielt daher auch für das Management von IT-Investitionen eine entscheidende Rolle. Folglich ist eine Disziplinen-übergreifende Integration von IT-Investitions- und Finanzmanagement notwendig, um IT-spezifische und finanzielle Aspekte integriert berücksichtigen zu können [30]. Eine Kombination etablierter Ansätze verschiedener Disziplinen bietet die Möglichkeit neue Lösungen für unternehmerische Fragestellungen zu entwickeln – selbst ohne eine grundlegende Weiterentwicklung der Ansätze. Gleichzeitig stellen wird eine wertvolle Grundlage geschaffen für weiterführende Arbeiten wie bspw. empirische Studien oder exemplarische Anwendungen und Weiterentwicklungen auf Basis von realen Daten.

Zur Bewertung der ökonomischen Potentiale von IT-Investitionen existieren zahlreiche, überwiegend empirische, Arbeiten (bspw. [4], [27], [38], [40]), die den Zusammenhang zwischen IT-Investitionen und unternehmerischem Erfolg untersuchen. Daneben existiert in der Literatur eine weitere Forschungsrichtung, in der auf Basis quantitativer bzw. finanzwirtschaftlicher Ansätze Methoden für eine risikointegrierte Bewertung von IT-Investitionen entwickelt und analysiert werden. Zur Berücksichtigung von IT-spezifischen Risiken erweitert bspw. [39] den allgemeinen Kapitalkostensatz um einen spezifischen IT-Risikozuschlag (WACC zu WACIT). [42] vermindern hingegen den stochastischen Kapitalwert der IT-Investitionen um einen IT-spezifischen Risikoabschlag. [41] und [21] betrachten in ihrer Arbeit insbesondere das Risiko durch stochastische Abhängigkeiten der IT-Investitionen.

In dieser Arbeit werden Ansätze des Investitions- und Finanzmanagements zur Berücksichtigung finanzieller Beschränkungen ([2], [13]) auf den Kontext der IT-Investitionsplanung übertragen und um eine entsprechende Risikobetrachtung erweitert. Dabei ist es nicht originäres Ziel diese Ansätze weiter zu entwickeln. Vielmehr sollen bestehenden Ansätze auf eine Problemstellung in einem anderen Kontext übertragen und entsprechend angepasst werden, um neue Lösungsmöglichkeiten zu entwickeln. Dadurch ist es möglich, finanzielle Rahmenbedingungen bei der Bewertung und Planung von langfristigen IT-Investitionen zu berücksichtigen. Mit Hilfe des entwickelten Optimierungsmodells kann gezeigt werden, dass der erwartete Wertbeitrag abgesichert bzw. sogar erhöht werden kann, wenn auf finanzielle Beschränkungen mit einer Liquiditäts- bzw. Budgetreserve – welche in der Finanz- bzw. Investitionsliteratur als probates Mittel anerkannt ist ([5], [18]) – reagiert wird. Um die Auswirkungen einer solchen Reserve auf den Wertbeitrag des IT-Investitionsvorhabens zu quantifizieren, wird dieser *mit* und *ohne* Reserve ermittelt und werden beide Alternativen miteinander verglichen.

Dazu wird ein Modell zur Planung von IT-Investitionen entwickelt, mit dessen Hilfe für ein Unternehmen die in Bezug auf Investitionszeitpunkt und Höhe optimalen Auszahlungen in das IT-Investitionsvorhaben unter Berücksichtigung finanzieller Rahmenbedingungen bestimmt werden können. Da es sich um einen integrierten Ansatz handelt, der IT-Investitions- und Finanzmanagement verbindet, wird die zur Entwicklung des Modells relevante Literatur an den jeweiligen Stellen herangezogen, sodass die Literatureinbettung im Modellteil und nicht wie zumeist üblich in einem gesonderten Literaturkapitel erfolgt.

Danach werden grundlegende Einflüsse zentraler Faktoren auf den Wertbeitrag des Investitionsvorhabens diskutiert und auf Basis eines Fallbeispiels analysiert. Zuletzt werden die Ergebnisse und Limitationen des Beitrags zusammengefasst und ein Ausblick auf weiterführende Forschungsfragen und praxisrelevante Implikationen gegeben.

2 Modell und relevante Literatur

Zunächst werden die Annahmen für die Bewertung des IT-Investitionsvorhabens anhand der relevanten Literatur vorgestellt und darauf aufbauend für das betrachtete Unternehmen die optimalen Investitionsauszahlungen (bzgl. Zeitpunkt und Höhe) zur Erreichung des maximalen Wertbeitrags (WB) ermittelt. Um den Einfluss der finanziellen Rahmenbedingungen auf den WB zu verdeutlichen, wird darauffolgend der resultierende WB in deren Abhängigkeit analysiert.

2.1 Modellannahmen

Im Rahmen eines Zwei-Perioden-Modells wird ein Unternehmen betrachtet, das die für ein mehrperiodiges IT-Investitionsvorhaben verfügbaren finanziellen Eigen- und Fremdmittel mit Hilfe einer Reserve optimal auf zwei mögliche Investitionszeitpunkte aufteilen möchte (vgl. Abbildung 1).

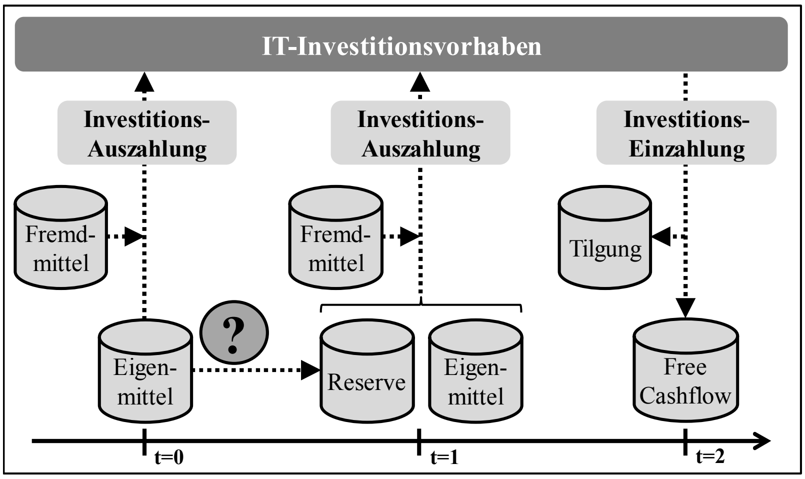


Abbildung 1. Entscheidungssituation des Modells

Annahme 1) *Das Unternehmen verfügt in den Zeitpunkten $t = 0$ und $t = 1$ über die Cashflows $\widehat{CF}_t$ aus dem bestehenden Unternehmensportfolio, die für die IT-Investitionen verwendet werden können. $\widehat{CF}_0$ ist bekannt, wohingegen $\widehat{CF}_1$ unbekannt und gemäß $\widehat{CF}_1 \sim N\left(E[\widehat{CF}_1], \sigma^2(\widehat{CF}_1)\right)$ normalverteilt ist.*

Bei der Planung mehrperiodiger IT-Investitionsvorhaben ist es nicht ausreichend, nur für den gegenwärtigen Zeitpunkt die Investitionsauszahlungen zu planen, da deren

Finanzierung über den kompletten Planungszeitraum hinweg sichergestellt sein muss. Andernfalls wir der insgesamt erwartete WB gefährdet, falls das Unternehmen zukünftige Investitionsauszahlungen nicht in ausreichender Höhe finanzieren kann.

Die Cashflows des Unternehmens stellen die Basis für die Planung der Investitionen dar. Dabei wird der zeitliche Horizont der Investitionen berücksichtigt, wobei sich Teilinvestitionen eines IT-Investitionsvorhabens in lang- und kurzfristige IT-Investitionen unterscheiden lassen (vgl. [24], [31], [43]). Dabei konkurrieren die Teilinvestitionen eines Investitionsvorhabens indirekt um begrenzte Unternehmensressourcen, wodurch sie sich gegenseitig beeinflussen und daher im Rahmen der IT-Investitionsplanung integriert betrachtet werden sollten [30].

Zur Finanzierung des Investitionsvorhabens kann das Unternehmen zusätzlich zu den Cashflows $\widehat{CF}_t$ auch Kredite aufnehmen und die für die Investition zu tätigenden Auszahlung I_t mit $t \in \{0,1\}$ dadurch erhöhen.

Annahme 2 a) *Im Zeitpunkt* t *nimmt das Unternehmen einen Kredit in Höhe von* K_t *mit* $t \in \{0,1\}$ *auf, welcher im Zeitpunkt* $t = 2$ *vollständig getilgt wird. Die mit dem risikolosen Zinssatz* r_f *anfallenden Zinsen sind zusammen mit der Tilgung im Zeitpunkt* $t = 2$ *endfällig zu zahlen.*

Insbesondere bei IT-Investitionen ist die Möglichkeit der Fremdfinanzierung allerdings oftmals stark eingeschränkt, da IT-Investitionen (bspw. die Entwicklung einer Software) i. d. R. zu einem hohen Anteil immateriell sind [22] bzw. mit Hilfe von IT-Investitionen unternehmensindividuelle, immaterielle Güter (bspw. Informationsverfügbarkeit und -qualität) geschaffen werden [10]. IT-Investitionen haben daher häufig einen spezifischen Wert für das jeweilige Unternehmen, der auch erhebliche immaterielle Investitionsbestandteile (bspw. aufgebautes Wissen) umfasst. Dieser unternehmensspezifische Wert ist durch den unternehmensspezifischen Wertanteil i. d. R. deutlich höher als ein rein anhand der Investitionsauszahlungen (I_t) feststellbarer objektiver Liquidationswert [11]. Darüber hinaus können IT-Investitionen bei einer kurzfristig notwendigen Liquidation nur mit erheblichen Preisabschlägen veräußert werden (falls überhaupt), da für ein anderes Unternehmen erneut erhebliche unternehmensspezifische Anpassungen erforderlich wären. Dadurch kann kurzfristig oftmals selbst der objektive Liquidationswert nicht erlöst werden.

Finanzielle Restriktionen dieser Art führen grundsätzlich zu einer Einschränkung der Investitionsfähigkeit eines Unternehmens [7], [12], [13]. Dies gilt insbesondere für IT-Investitionen, da deren hoher immaterieller Werteanteil die Aufnahme von Krediten zur Finanzierung der Investitionsauszahlungen erschwert [1], [20], [33]. Diese Besonderheit muss daher bei der IT-Investitionsplanung berücksichtigt werden, da dadurch die maximal finanzierbaren Investitionsauszahlungen begrenzt werden [1], [20], [33].

Annahme 2 b) *Die Höhe des Kredits* K_t *im Zeitpunkt* $t \in \{0,1\}$ *ist gemäß* $K_t \leq (1 - \gamma_{I_t}) \cdot I_t$ *beschränkt und hängt von der Höhe der Investitionsauszahlungen* I_t *des Unternehmens im jeweiligen Zeitpunkt ab. Dabei bezeichnet* $\gamma_{I_t} \in (0,1)$ *einen IT-investitionsspezifischen Illiquiditätsparameter.*

Der Illiquiditätsparameter γ_{I_t} bildet die beschränkte Fremdfinanzierungsmöglichkeit von IT-Investitionen ab. Die Höhe des Illiquiditätsparameters ist investitionsspezifisch, wobei bspw. ein erworbener Server (als standardisierte Hardware leichter übertragbar) i. d. R. mit einem geringeren Abschlag liquidiert werden kann als eine speziell entwickelte Individualsoftware (i. d. R. nur nutzbar für das jeweilige Unternehmen).

Um die verfügbaren Mittel optimal auf die periodischen Investitionsauszahlungen allokieren zu können, muss ein Unternehmen schätzen, wie die Höhe der erreichbaren Investitionseinzahlungen (d.h. die Rückflüsse der Investitionen) von der Höhe der Investitionsauszahlungen abhängt. Dadurch kann unter Berücksichtigung der geschätzten Grenzein- bzw. -auszahlungen der Wertbeitrag des IT-Investitionsvorhabens maximiert werden. Dazu muss ein Unternehmen das ökonomische Potential, d.h. die für das Unternehmen realisierbaren Einzahlungen, des IT-Investitionsvorhabens schätzen. Dabei ist zu berücksichtigen, dass verschiedene IT-Investitionen teilweise deutlich unterschiedliche ökonomische Potenziale aufweisen, die auch vom Investitionszeitpunkt und –umfang beeinflusst werden können. So weist bspw. eine frühzeitige Investition in eine IT-Innovation i.d.R. ein sehr hohes ökonomisches Potential auf, da sich dadurch bspw. neue Geschäftsfelder und Märkte erschließen lassen. Eine Investition in eine neue Verwaltungssoftware hat dagegen i.d.R. ein niedrigeres ökonomisches Potential, da sich damit zwar Effizienzgewinne erzielen lassen, diese aber oftmals deutlich geringer sind.

Zur Schätzung des Verhältnisses von monetären und nichtmonetären Inputgrößen zu Outputgrößen skalierbarer IT-Investitionen kann bspw. eine Funktion auf Grundlage der allgemeinen Produktionstheorie [37] verwendet werden (bspw. [8], [9], [16]). Solche Produktionsfunktionen finden bei der Planung von IT-Investitionen Anwendung [3], [22], wobei der dabei angenommene Zusammenhang i. d. R. nicht linear ist. Insbesondere bei großen IT-Investitionen sinkt mit steigendem Input der marginale Output, da mit steigendem Umfang die Komplexität und der Organisationsaufwand überproportional zunehmen [6], [22] , d.h. die Einzahlungen sind oftmals durch abnehmende Grenzeinzahlungen gekennzeichnet.

Darüber hinaus lassen sich aufgrund der komplexen und vielfältigen Abhängigkeitsbeziehungen und Wechselwirkungen die Ein- und Auszahlungen von IT-Investitionen ex ante nur schwer abschätzen, weshalb diese mit einer großen Unsicherheit behaftet sind [16], [22]. Dabei tragen insbesondere die unsicheren Einzahlungen einer IT-Investition zu deren finanziellen Risiko bei [29], [34]. Zur Vermeidung von Fehlallokationen des IT-Budgets ist daher das Risiko bzw. die Unsicherheit der Einzahlungen von IT-Investitionen im Rahmen der IT-Investitionsplanung zu berücksichtigen [24].

Annahme 3) *Die im Zeitpunkt* $t = 2$ *durch die Investitionsauszahlungen für das Unternehmen ermöglichten Einzahlungsüberschüsse werden durch die Funktionen* $EZ_i(I_t) = k_i \cdot f_i(I_t)$ *mit* $k_i \sim N(E[k_i], \sigma^2(k_i))$ *mit* $i = \{0,1\}$ *abgebildet. Dabei bildet* k_i *das unsichere ökonomische Potenzial der IT-Investition ab und* $f_i(I_t)$ *sind streng monoton steigende und konkave Funktionen. Diese repräsentieren für das Unternehmen den Barwert aller in Abhängigkeit von den Investitionsauszahlungen* I_t *in der Zukunft realisierbaren Einzahlungsüberschüsse.*

Wie bereits beschrieben verringert die Unsicherheit über zukünftig verfügbare Eigen- und Fremdmittel die Investitionsfähigkeit eines Unternehmens [7], [12], [13]. Zur finanziellen Absicherung existieren verschiedene Methoden, wobei in dieser Arbeit das Vorhalten einer Reserve betrachtet wird, da viele Unternehmen bspw. nach [5], [18] hohe Summen als finanzielle Reserven vorhalten. Darüber hinaus ist bspw. eine Betrachtung von Kreditlinien im Rahmen der IT-Investitionsplanung nicht sinnvoll, da IT-Investitionen durch ihre hohe Illiquidität nur sehr eingeschränkt als Sicherheit verwendet werden können und mit Kreditlinien weitere Beschränkungen verbunden sind [12].

Annahme 4) *Das Unternehmen bildet in* $t = 0$ *eine (unverzinsliche) Reserve* $R \geq 0$, *welche in* $t = 1$ *zusätzlich für die Finanzierung der Investitionsauszahlung zur Verfügung steht.*

Das Unternehmen steht somit vor der Entscheidung, in welcher Höhe die aktuell verfügbaren finanziellen Mittel sofort investiert oder teilweise als Reserve für eine Absicherung der Investition in $t = 1$ vorgehalten werden sollen. Auf Basis der getroffenen Annahmen ergeben sich für das Unternehmen die folgenden Cashflows CF_t für $t \in \{0,1,2\}$, die die Grundlage für die Entscheidung bilden:

$$CF_0 = \widehat{CF}_0 + K_0 - I_0 - R$$

$$CF_1 = \widehat{CF}_1 + K_1 - I_1 + R$$

$$CF_2 = EZ_0(I_0) + EZ_1(\tilde{I}_1) - K_0 \cdot \left(1 + r_f\right)^2 - K_1 \cdot (1 + r_f) \tag{1}$$

Dabei bezeichnet $CF_t \geq 0$ den nachschüssig anfallenden Free Cashflow, der für Ausschüttungen an die Eigenkapitalgeber zur Verfügung steht und die Grundlage für eine objektive Bewertung der IT-Investition bildet [19].

Um den Zeitwert der Zahlungen zu berücksichtigen, wird der Wert des Investitionsvorhabens zum Bewertungszeitpunkt $t = 0$ mit Hilfe eines Discounted Cashflow-Verfahrens ermittelt [25]. Die Risiken der unsicheren Zahlungsüberschüsse können dabei mit Hilfe der Risikozuschlags- oder Risikoabschlagsmethode berücksichtigt werden, welche bei konsistenter Anwendung zu identischen Ergebnissen führen [14], [32]. Zur Bewertung mehrperiodiger Investitionen kann die Risikoabschlagsmethode verwendet werden [28], bei der ein Abzug eines Risikoabschlags von den erwarteten Free Cashflows gemäß $E[CF_t] - RA(CF_t)$ erfolgt. Der Risikoabschlag $RA(CF_t)$ wird dabei basierend auf dem Capital Asset Pricing

Model (CAPM)[1] [17] multiplikativ mit Hilfe des Marktpreises des Risikos $\lambda = \frac{E[r_M]-r_f}{\sigma^2(r_M)}$ und der Kovarianz des Zahlungsüberschusses CF_t mit der Rendite r_M des Marktportfolios[2] $Cov(CF_t, r_M)$ ermittelt [23], [36]:

$$RA(CF_t) = \lambda \cdot Cov(CF_t, r_M) \tag{2}$$

Somit kann ein risikoadjustierter WB für den Zahlungsstrom der IT-Investition als Summe der auf den Zeitpunkt $t = 0$ mit dem risikolosen Zinssatz r_f diskontierten Cashflows abzüglich der jeweiligen Risikoabschläge berechnet werden. Dabei soll der erwartete WB durch die Wahl der optimalen Höhe der Investitionsauszahlungen I_t^* mit $t \in \{0,1\}$ maximiert werden, wodurch sich das folgende Optimierungsproblem ergibt:

$$\max_{I_t} WB = \sum_{t=0}^{2} \frac{E[CF_t]-RA(CF_t)}{(1+r_f)^t} = CF_0 + \frac{E[CF_1]-\lambda\cdot Cov(CF_1,r_M)}{1+r_f} + \frac{E[CF_2]-\lambda\cdot Cov(CF_2,r_M)}{(1+r_f)^2}$$

$$CF_t \geq 0 \;\; \textit{für}\; t \in \{0,1,2\}$$

$$K_t \leq \left(1-\gamma_{I_t}\right) \cdot I_t \;\; \textit{für}\; t \in \{0,1\} \tag{3}$$

Da im Optimum der Barwert der erwarteten Grenzzahlungsüberschüsse aus dem IT-Investitionsvorhaben dem Barwert der marginalen Investitionsauszahlungen entsprechen muss, können die optimalen Investitionsauszahlungen I_t^* folgendermaßen bestimmt werden:

$$\frac{\partial WB}{\partial I_0} = 0 \Leftrightarrow \frac{E[EZ_0'(I_0^*)]-\lambda\cdot f_0'(I_0^*)\cdot Cov(k_0,r_M)}{(1+r_f)^2} = 1 \;\; \textit{für}\;\; t = 0 \tag{4}$$

$$\frac{\partial WB}{\partial I_1} = 0 \Leftrightarrow \frac{E[EZ_1'(I_1^*)]-\lambda\cdot f_1'(I_1^*)\cdot Cov(k_1,r_M)}{(1+r_f)^2} = \frac{1}{(1+r_f)} \;\; \textit{für}\;\; t = 1 \tag{5}$$

Von I_t^* abweichende Investitionsauszahlungen vermindern den Wertbeitrag, da einerseits höhere Investitionsauszahlungen $(I_t > I_t^*)$ nicht durch die hieraus zusätzlich resultierenden risikoadjustierten Zahlungsüberschüsse kompensiert werden und andererseits niedrigere Investitionsauszahlungen $(I_t < I_t^*)$ weitere, noch mögliche Steigerungen des Wertbeitrags nicht realisieren.

2.2 Analyse des Modells

Im Folgenden wird die Investitionsstrategie eines Unternehmens analysiert, das finanziellen Beschränkungen unterliegt, d.h. es kann die eigentlich optimalen Investitionsauszahlungen I_t^* auf Grund von zu niedrigen bzw. zu unsicheren Cashflows $\widehat{CF}_t$ oder zu hohen Illiquiditätsabschlägen γ_{I_t} bei der Kreditaufnahme nicht finanzieren.

---

[1] Es liegen keine Unsicherheit und intertemporalen Abhängigkeiten bzgl. der Verteilungsparameter vor, sodass eine mehrperiodige Anwendung des CAPM möglich ist.

[2] Das Marktportfolio enthält alle am Kapitalmarkt verfügbaren Vermögenswerte im Verhältnis zu ihrer Kapitalisierung.

Somit kann das Unternehmen selbst bei Verwendung aller zu den Investitionszeitpunkten $t \in \{0,1\}$ verfügbaren Mittel nur die geringeren *Investitionsauszahlungen eines finanziell beschränkten Unternehmens* $I_t^B < I_t^*$ finanzieren. Zur Maximierung des WB kann das Unternehmen mit Hilfe einer Reserve die Investitionsauszahlungen I_t^B mit $t \in \{0,1\}$ folgendermaßen optimieren:

Einerseits wird die Investitionsauszahlung I_0^B durch die Bildung der Reserve verringert und andererseits die Investitionsauszahlung I_1^B erhöht. Folglich verringert bzw. erhöht sich der resultierende Zahlungsüberschuss der jeweiligen Teilinvestition des Investitionsvorhabens. Je nachdem, welche Veränderung die höheren Auswirkungen hat, wird durch die Einzahlung in eine Reserve der insgesamt aus dem Investitionsvorhaben resultierende Wertbeitrag negativ oder positiv beeinflusst. Zur Maximierung muss daher die optimale Höhe der Reserve R^* bestimmt werden.

Das Unternehmen wird in beiden Investitionszeitpunkten alle verfügbaren Mittel investieren, weil die Investition höhere Grenzein- als -auseinzahlungen erwarten lässt, da $I_t^B < I_t^*$ für $t \in \{0,1\}$. Zusätzlich werden die Investitionen mit so hohen Krediten wie möglich finanziert um die Investitionsauszahlungen zu erhöhen. Somit ergibt sich I_0^B mit Hilfe von Annahme 2b) und Formel (1) folgendermaßen:

$$I_0^B = \frac{\widehat{CF}_0 - R}{\gamma_{I_0}} \tag{6}$$

Sie ist von der Höhe der verfügbaren finanziellen Mittel sowie von der Höhe der gebildeten Reserve abhängig. Analog dazu ist in $t = 1$ die Höhe der Investitionsauszahlung I_1^B ebenfalls von der Höhe der risikoadjustierten finanziellen Mittel und der Höhe der in diesem Zeitpunkt aufgelösten Reserve gemäß

$$I_1^B = \frac{E[\widehat{CF}_1] - \lambda \cdot Cov(\widehat{CF}_1, r_M) + R}{\gamma_{I_1}} \tag{7}$$

abhängig. Da in beiden Investitionszeitpunkten alle verfügbaren finanziellen Mittel investiert werden ($CF_0 = CF_1 = 0$), ist nur der resultierende Cashflow des Zeitpunkts $t = 2$ für die Bestimmung der optimalen Höhe der Investitionsauszahlungen relevant. Das Optimierungsproblem des Unternehmens lässt sich damit nach

$$\max_{I_t} WB^B = \frac{E[k_0] \cdot f_0(I_0^B) - \lambda \cdot f_0(I_0^B) \cdot Cov(k_0, r_M)}{(1+r_f)^2} + \frac{E[k_1] \cdot f_1(I_1^B) - \lambda \cdot f_1(I_1^B) \cdot Cov(k_1, r_M)}{(1+r_f)^2}$$

$$- \frac{(1-\gamma_{I_0}) \cdot (I_0^B) \cdot (1+r_f)^2}{(1+r_f)^2} - \frac{(1-\gamma_{I_1}) \cdot (I_1^B) \cdot (1+r_f)}{(1+r_f)^2} \tag{8}$$

als Summe der risikoadjustierten Barwerte der erwarteten Zahlungsüberschüsse aus dem IT-Investitionsvorhaben abzüglich der Barwerte der Tilgung der aufgenommenen Kredite (inklusive Zinsen) abbilden.

Wie bereits erläutert hat die Reserve für das Unternehmen sowohl erhöhende, als auch verringernde Auswirkungen auf den Wertbeitrag. Daher werden so viele Mittel in die Reserve eingestellt, bis die marginale Erhöhung des WB (durch die höhere Investitionsauszahlung in $t = 1$) der marginalen Verringerung des Wertbeitrags (durch die geringere Investitionsauszahlung in $t = 0$) entspricht. Diese optimale Höhe der finanziellen Reserve R^* kann mit Hilfe der folgenden Bedingung bestimmt werden:

$$\left[E[k_0] \cdot f_0'(I_0^B) - \lambda \cdot f_0'(I_0^B) \cdot Cov(k_0, r_M) - \left(1 - \gamma_{I_0}\right) \cdot \left(1 + r_f\right)^2\right] \cdot \left(\frac{1}{\gamma_{I_0}}\right)$$

$$= \left[E[k_1] \cdot f_1'(I_1^B) - \lambda \cdot f_1'(I_1^B) \cdot Cov(k_1, r_M) - \left(1 - \gamma_{I_1}\right) \cdot \left(1 + r_f\right)\right] \cdot \left(\frac{1}{\gamma_{I_1}}\right) \Leftrightarrow$$

$$\Leftrightarrow \left[E[k_0] \cdot f_0'\left(\frac{\widehat{CF}_0 - R^*}{\gamma_{I_0}}\right) - \lambda \cdot f_0'\left(\frac{\widehat{CF}_0 - R^*}{\gamma_{I_0}}\right) \cdot Cov(k_0, r_M) - \left(1 - \gamma_{I_0}\right) \cdot \left(1 + r_f\right)^2\right] \cdot \left(\frac{1}{\gamma_{I_0}}\right)$$

$$= \begin{bmatrix} E[k_1] \cdot f_1'\left(\frac{E[\widehat{CF}_1] + R^* - \lambda \cdot Cov(\widehat{CF}_1, r_M)}{\gamma_{I_1}}\right) \\ -\lambda \cdot f_1'\left(\frac{E[\widehat{CF}_1] + R^* - \lambda \cdot Cov(\widehat{CF}_1, r_M)}{\gamma_{I_1}}\right) \cdot Cov(k_1, r_M) - \left(1 - \gamma_{I_1}\right) \cdot \left(1 + r_f\right) \end{bmatrix} \cdot \left(\frac{1}{\gamma_{I_1}}\right) \tag{9}$$

Unter Berücksichtigung der marginalen Erhöhung bzw. Verringerung des WB kann ein finanziell beschränktes Unternehmen durch optimale Wahl der Höhe der finanziellen Reserve R^* den resultierenden WB maximieren. Diese Möglichkeit zur Maximierung besteht allerdings nur für ein Unternehmen, welches finanzielle Beschränkungen bereits bei der IT-Investitionsplanung berücksichtigt.

Um die Auswirkung der Reserve auf den WB zu quantifizieren, wird für ein finanziell beschränktes Unternehmen der WB mit optimaler Höhe der Reserve $WB_{R^*}^B$ (d.h. unter Berücksichtigung der finanziellen Beschränkungen im Rahmen der IT-Investitionsplanung) mit dem WB ohne Reserve $WB_{R=0}^B$ verglichen. Somit kann eine Wertbeitragssteigerung WBS_{R^*} als Differenz der beiden Wertbeiträge definiert werden:

$$WBS_{R^*} = WB_{R^*}^B - WB_{R=0}^B \geq 0 \tag{10}$$

Im Folgenden wird das Modell auf Basis einer exemplarische Anwendung analysiert, um die Effekte geänderter finanzieller Beschränkungen zu quantifizieren.

3 Exemplarische Anwendung des Modells

Die folgende exemplarische Anwendung hat zum Ziel, zentrale Wirkungszusammenhänge des Modells mit Hilfe von Sensitivitätsanalysen zu verdeutlichen. Dabei werden auch die für ein Unternehmen relevanten ökonomischen Auswirkungen einer Reserve veranschaulicht. Für die exemplarische Anwendung werden zunächst die optimalen Investitionsauszahlungen ermittelt – ohne und mit Berücksichtigung der finanziellen Situation des Unternehmens – und anschließend mittels der Sensitivitätsanalysen der Einfluss relevanter Parameter auf den resultierenden WB bzw. die mögliche WBS_{R^*} analysiert.

Da eine Evaluierung des Modells auf Basis von Realweltdaten noch aussteht, wird im vorliegenden Beitrag als erster Schritt eine Analyse auf Basis exemplarischer Werte bzw. Funktionen vorgenommen. Dabei wird von den in Tabelle 1 aufgeführten Funktionen ausgegangen, sodass sich durch Abzug der Risikokosten von den

Einzahlungen und deren Optimierung anhand Formel (4) bzw. (5) optimale Investitionssauzahlungen I_t^* i. H. v. 55,47 bzw. 113,38 Geldeinheiten (GE) ergeben.

Tabelle 1. Einzahlungen, Risikokosten und optimale Investitionsauszahlungen

Zeitpunkt	$EZ_i(I_t)$[3]	$\lambda \cdot f_i(I_t) \cdot Cov(k_i, r_M)$	I_t^*
$t = 0$	$1 \cdot 100 \cdot ln(I_0)$	$2{,}67 \cdot 100 \cdot ln(I_0) \cdot 0{,}15$	55,47
$t = 1$	$1 \cdot 150 \cdot ln(I_1)$	$2{,}67 \cdot 150 \cdot ln(I_1) \cdot 0{,}075$	115,38

Über die zur Finanzierung der Investitionsauszahlungen verfügbaren Cashflows (d.h. Eigenmittel) liegen Schätzungen zur erwarteten Höhe und deren Risiko vor. Die eingeschränkte Kreditaufnahme ist durch den Illiquiditätsparameter γ_{I_t} abgebildet. Tabelle 2 enthält die entsprechenden Werte zur Höhe der in den Investitionszeitpunkten verfügbaren Eigen- und Fremdmittel.

Tabelle 2. Parameterwerte zur finanziellen Situation des Unternehmens

	$t = 0$	$t = 1$	$t = 2$
Bestehender Cashflow $\widehat{CF}_t$	20	20	20
Standardabweichung der Cashflows $\sigma(\widehat{CF}_t)$	-	10	15
Illiquiditätsparameter γ_{I_t}	0,5	0,9	-

Auf Basis der in den Investitionszeitpunkten beschränkten Eigen- und Fremdmittel plant das Unternehmen für den Zeitpunkt $t = 0$ mit 40 GE (vgl. Formel (6)) und für den Zeitpunkt $t = 1$ mit 15,55 GE (vgl. Formel (7)) als erwartete Mittel, die für das IT-Investitionsvorhaben zur Verfügung stehen. Mit diesen Werten, die deutlich unter den optimalen Investitionsauszahlungen liegen (vgl. Tabelle 1), kann nur ein erwarteter WB i. H. v. 497,79 GE erreicht werden (vgl. Formel (3)). Werden dagegen mit Hilfe einer Reserve verfügbare Mittel i. H. v. 9,96 GE von $t = 0$ auf $t = 1$ übertragen, so können Investitionsauszahlungen i. H. v. 20,05 GE in $t = 0$ und 26,2 GE in $t = 1$ finanziert werden, die ebenfalls deutlich unter den optimalen Investitionsauszahlungen liegen, aber einen höheren WB i. H. v. 528,07 GE (vgl. Formel (8)) ermöglichen.

Tabelle 3. Vergleich der Situation mit und ohne Reserve

	R^*	I_0^B	I_1^B	$WB_{R^*}^B$
Mit Reserve	9,96	20,05	26,2	528,07
Ohne Reserve	-	40	15,56	497,79

Ohne Reserve würde das Unternehmen in $t = 0$ einen zu hohen Betrag und in $t = 1$ einen zu niedrigen Betrag investieren. Die Reserve ermöglicht eine optimale Anpassung der Investitionsauszahlungen, wodurch eine Wertbeitragssteigerung

[3] Die Funktionen bilden den Barwert aller zukünftig anfallenden Einzahlungsüberschüsse zum Zeitpunkt $t = 2$ ab.

WBS_{R^*} i. H. v. 30,28 GE (~ 6,1%) erreicht werden kann. Um zu verdeutlichen, wie WBS_{R^*} von den finanziellen Rahmenbedingungen beeinflusst wird, werden die folgenden Sensitivitätsanalysen durchgeführt.

3.1 Analyse des Cashflows aus dem Unternehmensportfolio in $t =$

Ist in $t = 0$ ein im Vergleich zu den Startwerten des Beispiels höherer Cashflow $\widehat{CF}_0$ verfügbar, werden die höheren Mittel auf die Investitionsauszahlung in diesem Zeitpunkt und die Reserve aufgeteilt. Je höher dabei $\widehat{CF}_0$ ist, desto höher ist auch die Einstellung in die Reserve, da das Unternehmen durch die Investitionsauszahlungen in $t = 1$ höhere Grenzeinzahlungen erreichen kann. Ohne Reserve würde der höhere Zahlungsüberschuss für die Investitionsauszahlungen in $t = 0$ verwendet werden, wodurch aufgrund der niedrigeren Grenzeinzahlungen nur ein geringerer WB erzielt werden könnte. Zur Veranschaulichung sind in Abb. 2 sind WBS_{R^*} und R^* in Abhängigkeit vom $\widehat{CF}_0$ abgebildet.

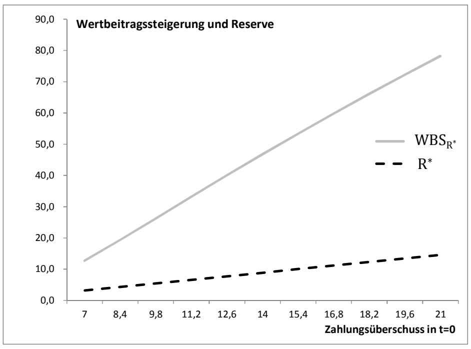


Abbildung 2. Einfluss der Höhe des Cashflows in $t = 0$

Der wertbeitragssteigernde Effekt der Reserve wird größer und eine höhere Reserve wird vorgehalten, je höher der in $t = 0$ verfügbare Cashflow $\widehat{CF}_0$ ist.

3.2 Analyse des Cashflows aus dem Unternehmensportfolio in $t = 1$

Erwartet das Unternehmen in $t = 1$ einen im Vergleich zu den Startwerten des Beispiels höheren Cashflow $\widehat{CF}_1$, kann die Höhe der Reserve verringert werden, da mehr Mittel zur Finanzierung der Investitionsauszahlung in $t = 1$ erwartet werden. Bei einem niedrigeren Cashflow $\widehat{CF}_1$ wird die Reserve dagegen erhöht, um die Investitionsfähigkeit in $t = 1$ abzusichern. Ohne Reserve kann das Unternehmen die verfügbaren Investitionsauszahlungen in $t = 1$ nicht erhöhen, obwohl eine höhere Investition vorteilhaft wäre, da sich in $t = 1$ höhere marginale Einzahlungen erreichen lassen. Die Wertbeitragssteigerung WBS_{R^*} wird daher auch überproportional größer, je niedriger der erwartete Cashflow $E[\widehat{CF}_1]$ ist (vgl. Abb. 3).

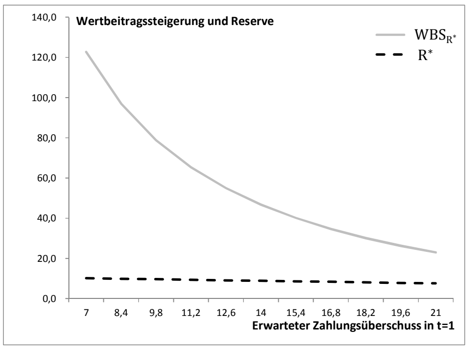


Abbildung 3. Einfluss der Höhe des erwarteten Cashflows in $t = 1$

Abb. 3 verdeutlicht, dass bei sinkendem erwarteten Cashflow $E[\widehat{CF}_1]$ durch eine Erhöhung der Reserve eine zunehmende Wertbeitragssteigerung WBS_{R^*} möglich ist.

3.3 Analyse des Risikos des Einzahlungsüberschusses in $t = 1$

Erhöht sich das Risiko des Cashflows $\widehat{CF}_1$ (d.h. $\sigma(\widehat{CF}_1)$) im Vergleich zu den Startwerten des Beispiels, sinken die verfügbaren risikoadjustierten Mittel in $t = 1$ durch die höheren Unsicherheit. Durch die Reserve kann dies bei der IT-Investitionsplanung berücksichtigt werden, sodass mehr Mittel für die Investition in $t = 1$ vorgehalten werden. Dabei steigt WBS_{R^*} wenn sich das Risiko des Cashflows $\sigma(\widehat{CF}_1)$ erhöht (vgl. Abb. 4).

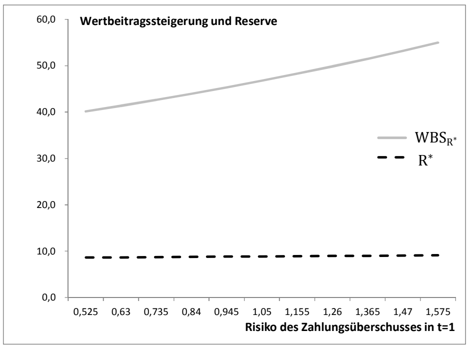


Abbildung 4. Einfluss des Risikos des Zahlungsüberschusses in $t = 1$

Insgesamt erhöht sich somit der Vorteil einer Reserve (gemessen durch WBS_{R^*}), je niedriger aktuelle und je höher zukünftige erwartete finanzielle Beschränkungen sind. Die Reserve hat somit einen hohen positiven Einfluss auf den WB von IT-Investitionen wenn ein Unternehmen gegenwärtig über hohe Mittel verfügt und davon einen Teil für Investitionen in zukünftigen Perioden vorhält, in denen die Finanzierung der geplanten Investitionsauszahlungen sehr unsicher ist.

4 Zusammenfassung und Ausblick

Durch finanzielle Beschränkungen besteht die Gefahr, dass langfristige IT-Investitionen eingeschränkt bzw. sogar kurzfristig abgebrochen werden müssen, wodurch ein Unternehmen deren geplanter WB nicht realisieren kann. Da die Relevanz von IT-Investitionen bspw. durch die voranschreitende Digitalisierung für Unternehmen verschiedenster Branchen weiter zunimmt und eine Berücksichtigung der finanziellen Rahmenbedingungen bei der Planung von solchen Investitionsvorhaben i. d. R. nicht ausreichend erfolgt, wurden in dieser Arbeit IT-Investitions- und Finanzmanagement in einen Ansatz integriert.

Mit dem entwickelten Modell können mit Hilfe einer finanziellen Reserve die negativen Auswirkungen finanzieller Beschränkungen verringert werden, da ein Unternehmen die Investitionsauszahlungen anpassen kann und dadurch der resultierende WB steigt. Zur Verdeutlichung des Einflusses der Reserve wurde die mögliche Wertbeitragssteigerung anhand eines Fallbeispiels analysiert. Dabei zeigen die durchgeführten Sensitivitätsanalysen, dass der wertbeitragssteigernde Effekt besonders deutlich ist, wenn ein Unternehmen aktuell über hohe Mittel verfügt und zukünftige Mittel geringer bzw. unsicher sind und dadurch die zukünftige Investitionsfähigkeit gefährdet ist. In diesem Fall kann durch die Reserve die langfristige Finanzierung des IT-Investitionsvorhabens abgesichert und dessen WB gesteigert werden.

Daher sollten Unternehmen bei der Budgetierung längerfristiger IT-Investitionsvorhaben nicht nur die Investitionsauszahlungen der aktuellen Periode optimieren, sondern auch zukünftige Investitionsauszahlungen berücksichtigen und deren Finanzierung absichern. Daher sollte eine Abstimmung zwischen Finanz- und IT-Abteilung erfolgen, damit die notwendigen Mittel zum richtigen Zeitpunkt zur Verfügung stehen. Dazu sollte ein langfristiges Optimierungskalkül zugrunde gelegt werden, bei dem neben IT-spezifischen Faktoren auch finanzielle Ertrags- und Risikoaspekte berücksichtigt werden. Auf dieser Basis können optimale langfristige Budgets bestimmt werden. Dabei sollte es einem IT-Projektmanager auch ermöglicht werden einen Teil des Budgets für Investitionsauszahlungen in Folgeperioden zu reservieren, ohne eine Budgetkürzung befürchten zu müssen.

Durch die getroffenen Annahmen und gewählten Parameter weist das Modell Schwächen bzw. Einschränkungen auf, die allerdings zugleich Raum für Erweiterungen im Rahmen zukünftiger Forschungsarbeiten bieten. So könnte für die Quantifizierung der Einzahlungen und Risiken Schätzverfahren entwickelt werden, die für verschiedene IT-Investitionen passende Funktionen auf Basis historischer Daten spezifizieren. Da der Fokus der Arbeit auf den finanziellen Aspekten der IT-Investitionen liegt, bleibt eine tiefergehende Analyse und Abbildung immaterieller Werte von IT-Investitionen künftiger Forschung vorbehalten. Da das auf bestehenden Ansätzen aufbaut und diese neu kombiniert, kann es keine fundamentale Weiterentwicklung dieser Ansätze leisten. Nichtsdestotrotz bietet das Modell einen Disziplinen-übergreifenden Ansatz, der im Rahmen weiterer Arbeiten empirisch validiert oder anhand eines realen Praxisbeispiels angewendet werden könnte (bspw. durch großzahlige empirische Untersuchungen oder Fallstudien).

Trotz der genannten Restriktionen kann der vorliegende Beitrag in einem ersten Schritt zeigen, dass eine Berücksichtigung finanzieller Rahmenbedingungen bei der Planung mehrperiodiger IT-Investitionen einen Beitrag zur Absicherung bzw. Erhöhung des resultierenden WB leisten kann.

Literatur

1. Almeida, H., Campello, M.: Financial Constraints, Asset Tangibility, and Corporate Investment. Rev. Financ. Stud., 20, pp. 1429-1460 (2007)
2. Almeida, H., Campello, M., Weisbach, M. S.: Corporate financial and investment policies when future financing is not frictionless. J. Corp. Fin., 17, pp. 675-693 (2011)
3. Banker, R. D., Chang, H., Kemerer, C. F.: Evidence on economies of scale in software development. Inf. Soft. Techn., 36, pp. 275-280 (1994)
4. Banker, R. D., Kauffman, R. J.: The evolution of research on on information systems - A Fiftieth-Year Survey of the Literature in Management Science. Manag. Sci., 50, pp. 281-298 (2004)
5. Bates, T. W., Kahle, K. M., Stulz, R. M.: Why do U.S. Firms Hold So Much More Cash than They Used To? J. Fin., 64, pp. 1985-2021 (2009)
6. Boehm, B., Clark, B., Horowitz, E., Westl, C., Madachy, R., Selby, R.: Cost models for the future life cycle processes: COCOMO 2. An. Softw. Eng., 1, pp. 57-94 (1995)
7. Boyle, G. W., Guthrie, G. A.: Investment, Uncertainty, and Liquidity. J. Fin., 58, pp. 2143-2166 (2003)
8. Brynjolfsson, E., Hitt, L.: Information Technology As A Factor Of Production: The Role Of Differences Among Firms. Eco. Innov. New Tech., 3, pp. 183-200 (1995)
9. Brynjolfsson, E., Hitt, L. M.: Computing Productivity: Firm-Level Evidence. Rev. Econ. Stat., 85, pp. 793-808 (2003)
10. Brynjolfsson, E., Hitt, L. M., Yang, S.: Intangible Assets: Computers and Organizational Capital. Brookings Papers on Economic Activity, 2002, pp. 137-181 (2002)
11. Brynjolfsson, E., Yang, S.: The Intangible Benefits and Costs of Investments: Evidence from Financial Markets. ICIS 1997 Proc., Paper 10, pp. 147-166 (1997)
12. Campello, M., Giambona, E., Graham, J. R., Harvey, C. R.: Liquidity Management and Corporate Investment During a Financial Crisis. Rev. Financ. Stud., pp. 1944-1979 (2011)
13. Cleary, S.: The Relationship between Firm Investment and Financial Status. J. Fin., 54, pp. 673-692 (1999)
14. Copeland, T. E., Weston, J. F., Shastri, K.: Financial Theory and Corporate Policy. 4th. Pearson Education Inc., Boston, MA (2005)
15. Dedrick, J., Gurbaxani, V., Kraemer, K. L.: Information technology and economic performance: A critical review of the empirical evidence. ACM Comp. Surv., 35, pp. 1-28 (2003)
16. Dewan, S., Shi, C., Gurbaxani, V.: Investigating the Risk-Return Relationship of Information Technology Investment: Firm-Level Empirical Analysis. Manag. Sci., 53, pp. 1829-1842 (2007)
17. Fama, E. F.: Risk-adjusted discount rates and capital budgeting under uncertainty. J. Financ. Econ., 5, pp. 3-24 (1977)
18. Ferreira, M. A., Vilela, A. S.: Why Do Firms Hold Cash? Evidence from EMU Countries. Eur. Financ. Manag., 10, pp. 295-319 (2004)
19. Franke, G., Hax, H.: Finanzwirtschaft des Unternehmens und Kapitalmarkt. 6. Springer, Berlin (2009)

20. Gamba, A., Triantis, A.: The Value of Financial Flexibility. J. Fin., 63, pp. 2263-2296 (2008)
21. Häckel, B., Hänsch, F.: Managing an IT Portfolio on a Synchronized Level or: The Costs of Partly Synchronized Investment Valuation. Dec. Sci., 23, pp. 388-412 (2014)
22. Krcmar, H.: Informationsmanagement. 5. Springer, Berlin (2010)
23. Kruschwitz, L., Husmann, S.: Finanzierung und Investition. 6th. Oldenbourg, München (2010)
24. Maizlish, B., Handler, R.: IT (Information Technology) Portfolio Management Step-by-Step: Unlocking the Business Value of Technology. 1st. Wiley, Hoboken (2005)
25. Perridon, L., Steiner, M., Rathgeber, A.: Finanzwirtschaft der Unternehmung. 15th. Vahlen, Munich (2009)
26. Gartner Says EMEA Enterprise IT Spending in Euros Will Decline 1.4 Percent in 2011 and Grow Only 2.3 Percent in 2012, http://www.gartner.com/it/page.jsp?id=1841115
27. Potthof, I.: Empirische Studien zum wirtschaftlichen Erfolg der Informationsverarbeitung. WIRTSCHAFTSINF., 40, pp. 54-65 (1998)
28. Röder, K., Müller, S.: Mehrperiodige Anwendung des CAPM im Rahmen von DCF-Verfahren. FB., pp. 225-233 (2001)
29. Ross, J. W., Beath, C. M.: Beyond the business case: New approaches to IT investment. MIT Sloan Manag. Rev., 43, pp. 51-59 (2002)
30. Santhanam, R., Kyparisis, G. J.: A decision model for interdependent information system project selection. Eur. J. Oper. Res., 89, pp. 380-399 (1996)
31. Schwartz, E. S., Zozaya-Gorostiza, C.: Investment under Uncertainty in Information Technology: Acquisition and Development Projects. Manag. Sci., 49, pp. 57-70 (2003)
32. Schwetzler, B.: Unternehmensbewertung unter Unsicherheit - Sicherheitsäquivalent- oder Risikozuschlagsmethode? Z. betriebswirt. Forsch., 52, pp. 469-486 (2000)
33. Shleifer, A., Vishny, R. W.: Liquidation Values and Debt Capacity: A Market Equilibrium Approach. J. Fin., 47, pp. 1343-1366 (1992)
34. Modernization: Clearing a Pathway to Success, http://www.standishgroup.com/newsroom/modernization.php
35. Trends in IT Investments, http://www.standishgroup.com/newsroom/investments.php
36. Steiner, M., Bruns, C.: Wertpapiermanagement. 9th. Schäffer-Poeschel Verlag, Stuttgart (2007)
37. Stiglitz, J. E.: Economics. W W Norton & Company, New York, London (1993)
38. Urbach, N., Smolnik, S., Riempp, G.: Der Stand der Forschung zur Erfolgsmessung von Informationssystemen - Eine Analyse vorhandener mehrdimensionaler Ansätze. WIRTSCHAFTSINF., 54, pp. 363-375 (2009)
39. Verhoef, C.: Quantifying the value of IT-investments. Sci. Comp. Prog., 56, pp. 315-342 (2005)
40. Walter, S. G., Spitta, T.: Approaches to the Ex-ante Evaluation of Investments into Information Systems. WIRTSCHAFTSINF., 46, pp. 171-180 (2004)
41. Wehrmann, A., Heinrich, B., Seifert, F.: Quantitatives IT-Portfoliomanagement: Risiken von IT-Investitionen wertorientiert steuern. WIRTSCHAFTSINF., 48, pp. 234-245 (2006)
42. Wehrmann, A., Zimmermann, S.: Integrierte Ex-ante-Rendite-/ Risikobewertung von IT-Investitionen. WIRTSCHAFTSINF., 47, pp. 247-257 (2005)
43. Weill, P., Broadbent, M.: Leveraging the New Infrastructure: How Market Leaders Capitalize on Information Technology. Harvard Business School Press (1998)

The Manifold Fruits of Digitalization – Determining the Literal Value Behind

Anna Neumeier[1], Thomas Wolf[1], and Severin Oesterle[2]

[1] Research Center Finance & Information Management, Augsburg, Germany
University of Augsburg
{anna.neumeier,thomas.wolf}@fim-rc.de
[2] Research Center Finance & Information Management, Bayreuth, Germany
University of Bayreuth
{severin.oesterle}@fim-rc.de

Abstract. As digitalization is rewriting the rules of competition, companies need to adapt to external changes or they will be left behind. Indeed, digitalization bears a lot of economic potential and is undoubted to have tremendous impacts on the economy. Many companies already launched digital initiatives. However, most of them lack an understanding of the value digitalization can create. They often neither know the organizational value created, nor define accountability measures or specify targets. Therefore, as a first step this paper aims to provide clarity by relating digital benefits listed in literature and highlighting the underlying value drivers. Our results help companies to identify digital business value, but also lower the hurdles that prevent them from scaling up their digital effort.

Keywords: Digitalization, benefit, value driver.

1 Introduction

Digitalization is on the rise and "is rewriting the rules of competition" [1]. Those companies who are not able or willing to adapt will be left behind [1–3]. Especially the ability to handle the new challenges that come with digitalization will make the difference in future. The main differences compared to common IT usage lie in the faster speed of change, in the higher level of interconnectedness, and in the willingness of individuals to use technological devices [2]. Those phenomena already have an enormous impact on private lives and the society and will also tremendously impact business practices and success models. Hence, digitalization is considered as the fourth wave of industrialization, which will dramatically change the business world [4].

Many consultancies provide extensive studies about success stories, trends, challenges, strategies and benefits of digitalization [5–8]. The economic potential of the Internet of Things, for instance, is considered to be up to $ 1.9 trillion by 2020 [9]. And yet, the Internet of things is only one of the topics considered as digitalization. Cloud computing, big data, social networks, and many other novelties unfold even more prosperous properties. Many companies set up digital initiatives in order to benefit from the supposed fruits of digitalization. A study among 500 company executives by KPMG showed that about 43% of companies already conducted changes in their business

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Neumeier, A.; Wolf, T.; Oesterle, S. (2017): The Manifold Fruits of Digitalization - Determining the Literal Value Behind, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 484-498

model and that about 60% already made adjustments towards digital in their product and service portfolios and their administrative processes [6]. However, companies mostly lack a deeper understanding of the value digitalization can create within their organization and therefore are mostly not able to determine the gain their digital initiatives actually shall deliver. According to a survey among 850 C-level executives by Gottlieb and Willmott [10], only 7% of the respondents consider their organization to understand the exact value from digitalization. Furthermore, over 60% admit that their companies do neither have accountability measures nor specified targets for digital initiatives [10]. This emphasizes the necessity for shading light to manifold supposed fruits of digitalization.

Therefore, the goal of our research endeavor is to identify the value of digitalization within an organization. Therefore, we need to clarify the main value drivers of digitalization that can be used as a starting point for developing measures for the value of digitalization. Thus, we state the following research question:

What are the benefits and which aspects can be considered as underlying value drivers of digitalization?

Therefore, we analyze benefits of digitalization that have been identified in existing literature and derive to the underlying value drivers. We contribute to existing literature by analyzing and structuring existing literature to provide the basis for a detailed analysis of measurement approaches for digital value drivers. Furthermore, practitioners can benefit from our approach as we provide an identification of main benefits and propose value drivers to determine the gain of digital initiatives.

The remainder of the paper is organized as follows. After motivating our research and introducing the research question, we explain our research methodology. In section 2, we provide a structured literature review and synthesize the results of the literature review in section 3. In section 4, we derive conclusions, reveal limitations and give an outlook on future research.

2 Literature Review

This paper sticks to the well-established method for research synthesis of Cooper and Hedges [11] and adopts the changes towards the seven-step process of Cooper [12]. Accordingly, we formulate the scope including research objective and problem in the introduction section. In this section we outline the procedure of searching literature and define which information about each study is relevant to the previously described research objective. The selection of relevant studies based on their contribution and fit to the research problem is also illustrated. To integrate, condense and combine the results, we analyze the articles and develop an incorporating framework in section 3. Subsequently, the results are summarized and conclusions are drawn in the discussion section with regard to generalization and possible points of contact for further research. Finally, with this paper the results of the research are presented to public.

To ensure transparency and objectivity regarding the research sample selection, we identified relevant literature based on a keyword search of different data bases. As digitalization is a quite young research field, which is extensively discussed in research and practice, we included scientific literature as well as practice studies in our search.

Procedure: To identify relevant practice studies that cope with the value of digitalization initiatives, we draw on google search engine and used ‘digi*’ AND ‘value’ and ‘digi*’ AND ‘benefits’ as keywords. We focused on articles in English and German language. Furthermore, we searched white papers and studies from large management consultancies (e.g., McKinsey & Company, Accenture) and conducted a forward and backward search within the identified practice studies.

Hence, we gained a vast amount of results from our google search. Thus, we searched for publications by management consultancies and other white papers. After identifying the most relevant articles by screening the titles, we analyzed the introduction of the available data set. Thus, we identified 29 articles and studies that deal with the benefits and value of digitalization for companies.

To identify relevant scientific literature, we conducted a structured literature analysis based on established scientific databases that are relevant in the IS field. With new technologies and innovations like cloud computing, big data and many others establishing, the rules of competition have changed. But what is new with the phenomenon of digitalization compared to the usage of IT or digital technology that has been going on for years already? In this context, Gimpel and Röglinger [2] state, that “the new aspect of digitalization is not the usage of information technology per se but the speed of change and the world´s level of connectedness”. The miniaturization of increasingly powerful computing hardware leads to an omnipresence of technological devices. This is associated as Moore’s law and is considered to be the root cause or underlying driver of the digitalization phenomenon [2]. Additionally, this interpenetrating dissemination of IT combined with innovations like the internet enable all entities of an ecosystem to connect, which boosts an increase in value. Simultaneously to these technological developments, customers are increasingly willing to use technological devices. All in all, digitalization can be described as a situation in which different influence factors interact and thus the usage of digital technology is able to tremendously impact business practices and business models across almost all industries or even the whole ecosystem.

Hence, we searched for articles that cope with digitalization and the value that can be gained for a company due to digitalization. Thus, the keywords ‘digi*’ AND ‘benefits’ as well as ‘value’ in the title and the abstract of the articles were used. Furthermore, we limited the search to articles that were published after 2010, since we want to put new use and omnipresence of digital devices to the center of contemplation. With the release of the first iPhone in 2007 and the first iPad in 2010, there was an enormous increase of smart devices per person [13]. In 2010, the ratio of more than one smart device per person (calculated on the basis of the world’s population) was observed for the first time [13]. Hence, we chose 2010 as an adequate time limit for our analysis. Table 1 shows the combination of search terms in the selected databases.

Table 1. Selected Databases and Search Terms

Database	*Keywords*
AIS Electronic Library	‘digi*’ AND ‘benefits’, ‘digi*’ AND ‘value’, ‘digitalization’
EbscoHost	‘digi*’ AND ‘benefits’, ‘digi*’ AND ‘value’
IEEEXplore	‘digi*’ AND ‘benefits’, ‘digi*’ AND ‘value’

As digitalization is an extremely broad research field, many articles can be identified that somehow deal with digitalization. Thus, the scientific literature review generated a result of more than 300 articles with duplicates removed. To identify the articles that are relevant for our research topic, we first analyzed the titles of the articles whether they fit the focus of our manuscript to identify the value of digitalization for companies. All articles that obviously cope with another topic were excluded. In a second step, we analyzed the abstracts of the remaining articles and focused on the articles that deal with the value of digitalization for companies. After analyzing all articles, we gained a result of 27 articles which address the identification of the value of digitalization for companies. To complete the literature review, we conducted a forward and backward search recommended by Webster and Watson [14]. The identified articles were analyzed as previously described and if appropriate added to the selected sample. Finally, 31 articles were identified as relevant for further consideration.

Analyses: Concerning management and practice studies, beside almost all big players in the consulting industry, independent research institutes, as well as public institutions address digitalization in various studies. Even though many articles do not focus on a specific industry, most industries (e.g., media, banking, insurance, and manufacturing) are mentioned in the articles.

The influence of digitalization and the implications for companies is one of the major concerns in practice studies [4, 8–10]. Even though they focus on different aspects like the expectations of companies [10], influencing factors of digitalization on business and working processes or the effects of digitalization on a special industry, like production industry, the insurance industry, or German mid-sized companies they all agree that digitalization already has changed and will change many aspects of a company [6, 15, 16].

This digital transformation comes with several chances and challenges [2, 17, 18]. Therefore, Capgemini set up a study concerning the management practices and benefits of the digital transformation [19]. To accomplish successful digital transformation, organizational requirements need to change and classic organizational structures will no longer be adequate [20]. Thus, transformation management can decide about success or failure of a company in the digital age [2, 18, 21]. Reactions and requirements to overcome the new challenge are also widely discussed. Furthermore, the success factors and benefits of digitalization for companies are examined [1, 7, 18, 19]. Accenture conduct a survey among the Top 500 German companies and identify strategies that make the difference for the successful digital players [22].

Scientific literature on digitalization also discusses various topics. Currently, the focus of companies changes from developing an IT strategy towards integrating a ‘digital business strategy’ [3]. Digital business strategy sheds light on the possibilities and effects of digitalization on the firm. In general, it provides special attention on making explicit that the new digital technologies go beyond increasing efficiency and productivity towards competitive advantage and strategic differentiation [3]. Value creation is not limited to single products or services anymore. It expands opportunities to new dimensions that have not been accessible before [23]. Companies in future will not be successful because they once adopted a digital business strategy or business model. They will succeed due to flexibility in adapting to new circumstances and requirements [23]. As the modes of collaboration within and beyond ecosystems will change, companies need to adapt and increase flexibility [24, 25]. Most articles concerning digital business strategy develop theoretical models concerning the digital architecture, digital networks, platforms and processes [e.g. 23–27].

Besides digital business strategy, the combination of digital and physical components is discussed in literature. Yoo et al. [28] define the "carrying out of new combinations of digital and physical components to produce novel products" as digital innovation. According to Barrett et al. [29] "the economy [shifts] from a goods-based to a service-based economy" due to digitalization and digital innovation [30]. The integration of digital technologies in traditional products promises increased performance and experience [31]. Thus, the combination of digital and physical products is becoming a strategic topic for many companies as the gained flexibility creates novel market offerings and thus benefits [28, 30, 32]. To gain the value that comes with new market offerings, new ways of defining digital services are required as it differs from the traditional logic of services [33]. Furthermore, traditional innovation processes need to change due to the integration of digital technologies [31].

Besides discussing the influence of digital technologies on the strategy of a company, the integration of digital technologies in the operations of companies is a focus topic in literature. Matt et al. [34] claim that if the four transformational dimensions 'use of technologies', 'changes in value creation', 'structural changes', and 'financial aspects' are closely aligned, the potential benefits of digitalization can be realized. This "include[s] increases in sales or productivity, innovations in value creation, as well as novel forms of interaction with customers, among others" [34]. The integration of new technologies into operations can furthermore leverage higher organizational efficiency and thus increase savings in cost and time [35, 36].

Digitalization can furthermore help to improve the communication within and beyond the company [37–39]. By building innovation communities companies can enlarge their idea pools and thus gain tangible as well as intangible benefits [39]. Due to easy communication in corporate blogs or platform forums the performance of employees can increase due to knowledge spillovers [37, 38]. Nevertheless, technology needs to be aligned with the business goal to improve overall performance.

Besides literature that deals with digitalization in general, there are several studies that examine digitalization within specific industry sectors, like automotive industry, retail banking, media, and public sector [40–42]. Even though all articles discuss different aspects and industry specifics, they agree that new business models can create new value-added services [40]. Improved communication with the customer as well as smart data analysis can help to recognize customer needs [42] and thus improve customer experience and satisfaction [41]. Furthermore, due to virtualization of decision and operating processes and standardization, companies can achieve an increase of efficiency that leads to higher productivity and thus reduced costs [40, 41].

Although various topics have already been discussed, neither in scientific nor in practice literature we can find a structured overview or framework that can help companies to understand the benefits of digitalization. Thus, we will examine the benefits of digitalization and the underlying value drivers in more detail.

3 Deriving the Value of Digitalization

Despite many studies from business consultancies about implications and possible benefits of digitalization, this topic is still very vague for many companies. They often are not able to grasp this topic in terms of being able to set concrete organizational objectives. Therefore, this section condenses the benefits previously investigated from scientific and practice literature in order to derive the value drivers of digitalization.

To receive a more objective result, we aligned to a structured approach and discussed all results with other researchers. As a first step we carefully screened the 29 practice studies and the 31 scientific articles for benefits of digitalization. We gathered all mentioned aspects in a central database. After having identified a variety of benefits of digitalization, we clustered all benefits independently of each other to get a common understanding of the various terms and notions mentioned. Then we discussed the results within each other and with other researchers. This led us to a result of 38 kinds of benefits. We derived an appropriate wording for each kind of benefit. For example, customer centric, delight customers, customer behavior, customer loyalty, customer trust, citizen-centricity, customer satisfaction, client experience were subsumed to the benefit *customer experience*. Due to space restrictions, we only included the final wording.

Since the identified benefits of digitalization are diverse regarding granularity and sphere of impact, we categorize them to different levels of an information system. Since the different elements of an information system usually are organized in organizational layers [43–45], we refer to an established model by Buhl and Kaiser [45] for information and communication systems to allocate them to appropriate layers of organizational information systems. This model has been applied for categorization by other studies before [46, 47]. The model classifies an organization in four different intra-organizational layers and is based on the classification by Frank [48] and Winter and Fischer [49]: *Business Model*, *Business Processes*, *Application Systems & Services (Appl. Sys & Services)*, and *Infrastructure*. The layers are connected and increase in granularity from the lower to the upper ones. Furthermore, the model considers *Customers* as value driver of each organizational endeavor and supposes them as extra-organizational layer, which is however closely connected to the intra-organizational business model.

Infrastructure mainly contains hardware and networks of a company [49] and describes all kinds of technology that are used within a company. The benefits that can be observed on this layer are caused by technological advances due to digitalization. The *Appl. Sys & Services* layer includes all software, application, and data components [49]. As more and more services are offered online, companies need to deploy IT applications to accomplish their business processes. The benefits that can be achieved within this layer are on a technical operations level but can help the company to accomplish the overall goals. The *Business Processes* layer pictures the procedure of performing service delivery to achieve the desired business goals [49] and thus this level includes all benefits that result from improvements on an functional operations level due to digitalization. In contrary to the process layer, the *Business Model* depicts the organization from a strategic point of view [48, 49] and should be aligned with customer needs. Thus it connects the intra- and extra-organizational level. All strategic benefits that can be observed in literature are assigned to this layer. The *Customer* layer depicts an extra-organizational layer that refers to all interactions between the company and the customer. Thus benefits that result from the interaction with customers can be assigned to this layer. Table 2 gives an overview of the identified benefits. An overview of all references and a detailed assignment of references to the benefits can be found an extra online repository [50].

Table 2. Categorization of Digitalization Benefits

Customer	
relevance among customer	product and service quality
innovative products and services	customer experience
customer interaction convenience	customer tailored solution
drive customer behavior	customer conversion
Business Model	
enlarge customer pool	advance to new business fields
profitability	increased sales
increase returns	risk mitigation
expand to digital channels	cost reduction
competitive advantage	enable innovations
enhanced promotion	efficiency
new competitive business models	
Business Processes	
incr. productivity	process flexibility
reduced product time-to market	speed of service proposition
operational excellence	process automation
smart workflow integration	process improvement
gain external network synergies	
Appl. Sys & Services	
improved information base	use of customer data
new delivery model	use of internal data
knowledge management	customer insights
real-time information	
Infrastructure	
smart technologies	

Based on this investigation, we recognized in which areas digitalization particularly is supposed to lead to benefits. However, it still was not clear, how these benefits may pay off and how this can be appropriately measured. Furthermore, the benefits of digitalization are partially interlinked and mutually dependent. To depict the first idea of relationships between the identified benefits, we used deductive logic to relate the manifold vague digital benefits and to conclude on measureable value drivers of digitalization. In doing so, we gradually assigned the different kinds of benefits not only to the different intra- and extra-organizational layers, but also to particular levels of granularity. Therefore, we used logic trees on each organizational level which visualize the different levels of granularity. Benefits of similar granularity have been assigned to the same level and differing ones have been assigned to different levels. Furthermore, we incrementally connected benefits on a lower level of granularity (e.g., *speed of service proposition* or *real-time information*) to benefits on higher levels of granularity (e.g., *increased productivity* or *profitability*), since the ones on the lower level logically contribute to corresponding ones on the higher level. Consequently, we derived a model that depicts a top-down and bottom-up perspective on all identified extra- and intra-organizational benefits related to digitalization.

Since the benefit of digitalization is a very vague research topic with only few existing scientific research, the objective of this paper was to open the field of research by structuring and relating different benefits mentioned in literature in a first step, rather than to state one particular hypothesis that needs to be tested. By structuring the existing benefits with logic trees, we used a qualitative, deductive method to actually state several different hypothesis about the relationship of the single benefits. In doing so, we tried to minimize subjective judgment by conducting a discussion group of diverse scientists. We had a core team of three researchers who discussed their findings and suggestions in several rounds with up to 20 other researchers, including five professors and several PhD graduates. Nevertheless, continuation research is still necessary to proof the stated hypothesis. The resulting model for the benefits of digitalization including their linkages and corresponding areas is depicted in Figure 1.

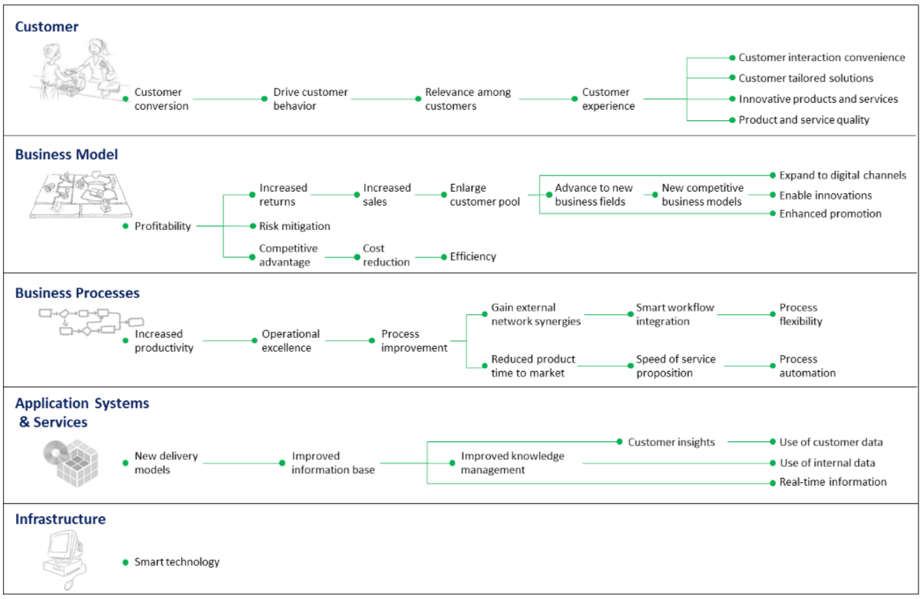


Figure 1. Logic Trees for the Digitalization Benefits on each Organizational Level

For all layers depicted in Figure 1, we exclusively used the benefits that were derived based on the analyzed literature and arranged them within logic trees. On the *Business Model* layer, Figure 1 shows the strategic benefits for an organization where, for instance, the *profitability* is influenced by three components: *increased returns*, *risk* *mitigation*, and *competitive advantage*. First, we observe higher *profitability* due to *increased returns* that can be caused by *increased sales*. Increased sales depend on an *enlarged customer pool* which can be generated by an *expansion to digital channels*, the *advance to new business fields*, and *enhanced promotion*. The advance to new business fields is enabled by *new competitive business models* that are based on digital *innovations.* Second, the *profitability* of a company can be influenced by risk. If improved *risk mitigation* strategies can be applied due to digitalization, profitability of a company can be increased. Third, profitability is influenced by the *competitive advantage* a company can achieve due to *cost reductions*, which in turn can be gained by improved *efficiency*. According to this example, the logic trees on the other organizational layers can be read similarly. Furthermore, the logic trees are not only connected within but also beyond the organizational layers.

4 Discussion

To enable a valid assessment of the target state of digitalization within a company, an appropriate measurement is required in future research. Therefore, we need to identify value drivers that can be measured to assess the status of digitalization. On the strategic level (*Business Model),* the overall profitability of the organization could be measured. As the overall profitability of a company is influenced by internal and external factors, one would not be able to obtain whether an increase in profitability would stem from an improved customer interaction (external factor) or an operational improvement (internal factor). Thus, we will not directly draw on the profitability to derive the overall contribution. However, as the *Business Model* layer constitutes the linkage between the intra- and the extra-organizational layer, the interfaces between Business Model & Customer as well as Business Model & Business Processes might offer a starting point for an appropriate measure. By measuring the consequences of digitalization at the intra- and the extra-organizational interfaces of the *Business Model* layer, an overall perspective on the status of digitalization can be gained. Hence, for each interface an appropriate value driver should be identified that (1) is able to depict the overall influence of digitalization on the intra- and extra-organizational layers and (2) can be influenced by the company. Thus, we choose the benefits on the highest granularity level that can be influenced by the company, i.e., as close to the root node of the logic trees as possible.

Intra-organizational perspective: The benefit on the highest granularity level from the intra-organizational perspective is competitive advantage. The competitive advantage of a company cannot easily be measured. Thus, cost reduction should be analyzed for its applicability as a measure for the intra-organizational perspective. Many companies use cost reductions to measure the success of initiatives that are supposed to increase internal competitiveness [51]. Nevertheless, cost reduction does not necessarily need to be the consequence of internal improvement. For example, production costs will decrease if less products are produced. In turn, less revenues can be achieved with less products to be sold, which will lead to an overall negative result. Thus, the sole measurement of cost reduction cannot cover particulars on success of digitalization. Thus, we examine *efficiency* as a value driver that can be measured. Efficiency is mainly driven by operational improvements of a company. In this paper, we consider efficiency from an economic perspective and hence suppose an activity to be efficient if there is no other way of performing a particular activity that leads to less cost, delivers the same results, upholds the same quality standards and is economically reasonable to invent. As efficiency, moreover, may be influenced by a company itself, we consider it to be an appropriate value driver for the intra-organizational perspective.

To assess efficiency within the organization established measures like maturity models or similar approaches can be used [52]. Thus, based on different maturity levels the status quo and target state of a company might be determined. A corresponding examination, however, is topic to further research.

Extra-organizational perspective: Within this perspective, there are many starting points for possible measures. For example, increased returns, increased sales, customer conversion and the size of the customer pool can quite easily be measured. However, those benefits are either not necessarily the consequence of external advancement or cannot be influenced directly by the company. Thus, we propose to choose *customer experience*, as this benefit is on the highest level of granularity that can be influenced by the company itself. While the other benefits like customer conversion oftentimes cannot be influenced from an internal point of view, customer experience can be directly changed with the help of appropriate projects.

A measure that might serve to determine a company's digital target state that maximizes the value driver customer experience is the Kano model [53]. This model measures customer experience through under and over fulfillment of customer expectations [54]. Based on a respectively designed questionnaire, a company could determine its customers' expectations, which might serve as a basis to determine the digital target state of the company. Another measure that could be used for assessing this perspective is the established American Customer Satisfaction Index [55]. A corresponding detailed examination is also in this case topic to further research.

Hence, *efficiency* and *customer experience* are identified as two of the main value drivers within digitalization that might be used as starting point to measure the value of digitalization and hence show the manifold fruits of digitalization. However, concrete measures as well as a reasonable process to assess the value of digitalization should be determined in future research.

5 Conclusion, Limitations and Outlook

As digitalization is on the rise, the competitive environment of companies in all different industries will change within the next years [1]. Thus, many companies set up digitalization initiatives to benefit from the promising prospects even though they lack a deeper understanding of the benefits digitalization can literally bring to their company. Furthermore, most companies are not able to assess their status quo as well as their target state concerning digitalization. To address the research question, we conducted a structured literature review within scientific and practice literature to identify the benefits of digitalization and to clarify the underlying value drivers. Thus, we identified 38 kinds of benefits of digitalization which can be categorized into the five areas of the model for organizational information and communication systems *Customer*, *Business Model*, *Business Processes, Application Systems & Services, and Infrastructure*. We used logic trees to relate the manifold vague digital benefits within each layer and to conclude on measureable value drivers of digitalization.

To enable an overall measurement, it is reasonable to consider the internal and the external perspective. Hence, we identified value drivers that enable the measurement at the interfaces between the intra- and extra-organizational perspectives. From the intra-organizational perspective efficiency enables to measure the influence of internal improvements. Measuring customer experience, furthermore, enables to measure digitalization progress from an external view. Each digitalization project should thus target to contribute to the company's profitability in terms of efficiency or customer experience or even both values. Thus, if a company decides to invest in digital initiatives, they should be assessed based on those drivers.

Despite the merits of this paper in terms of clarifying benefits and value drivers of digitalization, it is not without limitations. We base our literature review on scientific and practice literature. While scientific literature sticks to objective research goals and methods, practice literature might inhabit a bias as management consultancies want to position themselves as digitalization experts. Thus, they might exaggerate the positive effects and downplay the risks and challenges to gain new assignments. Nevertheless, management consultancies were able to gain deep insights into the real world application of digitalization during the last years. As the main focus of our research endeavor is to identify the key benefits and value of digitalization, we consider practice studies to be the most fertile source of information. Furthermore, we manually clustered and categorized the identified benefits to different kinds of benefits and layers of the organization. This might lead to subjectivity. However, we tried to reduce subjectivity as far as possible by drawing on a structured research procedure and by discussing our results with other researchers. Moreover, the developed categorization of benefits and the identified value drivers are a first approach to clarify the value of digitalization. However, we did not yet derive an empirical investigation of the underlying hypothesis. This is topic to further research. Furthermore, future research should feel encouraged to analyze different measurement methods for digital value drivers as well as establish a process for the valuation of digitalization projects.

Overall, our study contributes to existing knowledge and approaches the research question raised within this manuscript. Our manuscript can be equally beneficial for research and practice. On the one hand, the approach contributes to future research by analyzing and structuring existing literature and can thus provide the basis for a detailed analysis of measurement approaches for digital value drivers. Furthermore, our results support practice with the identification of the main benefits and value drivers of digitalization. Thus, they can make use of the categorization to determine the gain of their digital initiatives. Thus, all in all, the approach can help practitioners to get a better understanding of the benefits of digitalization, which might lower the hurdles that prevent companies from scaling up their digital effort.

6 Acknowledgements

Grateful acknowledgement is due to the DFG (German Research Foundation) for their support of the project „Value-based Management of IT Projects“ (FR 2987/2-1; BU 809/13-1) making this paper possible.

References

1. Hirt, M. and Willmott, P.: Strategic Principles for Competing in the Digital Age, http://www.mckinsey.com/insights/strategy/strategic_principles_for_competing_in_the_digital_age (2014) (Accessed: 18.08.2016)
2. Gimpel, H., Röglinger, M.: Digital Transformation: Changes and Chances. Project Group Business and Information Systems Engineering (BISE) of the Fraunhofer Institute for Applied Information Technology (FIT), Augsburg / Bayreuth (2015)
3. Bharadwaj, A., El Sawy, O.A., Pavlou, P.A., Venkatraman, N.: Digital Business Strategy: Toward a Next Generation of Insights. Management Information Systems Quarterly 37, 471–482 (2013)
4. Bloching, B. and Wege, E.: Wer teilt, gewinnt, http://www.rolandberger.de/media/pdf/Roland_Berger_TAB_Wer_teilt_gewinnt_20140718.pdf (2014) (Accessed: 17.07.2016)
5. Jaubert, M., Marcu, S., Ullrich, M., Malbate, J.-B. and Dela, R.: Going Digital: The Banking Tranformation Roadmap, https://www.atkearney.com/digital-business/ideas-insights/featured-article/-/asset_publisher/Su8nWSQlHtbB/content/going-digital-the-banking-transformation-road-map/10192 (2014) (Accessed: 17.07.2016)
6. Gutsche, R.: Survival of the Smartest - Which Companies Will Survive the Digital Revolution?, https://www.kpmg.com/DE/de/Documents/survival-of-the-smartest-2014-kpmg-en.pdf (2014) (Accessed: 01.08.2016)
7. Markovitch, S. and Willmott, P.: Accelerating the Digitalization of Business Processes, http://www.mckinsey.com/insights/business_technology/accelerating_the_digitization_of_business_processes (2014) (Accessed: 16.08.2016)
8. Olanrewaju, T. and Willmott, P.: Finding your digital sweet spot, http://www.mckinsey.com/insights/business_technology/finding_your_digital_sweet_spot (2013) (Accessed: 10.08.2016)
9. Gartner: The Nexus of Forces, http://www.gartner.com/technology/ research/nexus-of-forces/ (2013) (Accessed: 22.08.2016)
10. Gottlieb, J. and Willmott, P.: McKinsey Global Survey Results - The Digital Tipping Point, http://www.mckinsey.com/insights/business_technology/ the_digital_tipping_point_mckinsey_global_survey_results (2014) (Accessed: 03.08.2016)
11. Cooper, H., Hedges, L.V.: The Handbook of Research Synthesis. Russell Sage Foundation, New York (1994)
12. Cooper, H.: Research Synthesis and Meta-Analysis: A Step-by-Step Approach. Sage publications, California (2010)
13. Cisco IBSG: The Internet of Things. How the Next Evolution of the Internet Is Changing Everything, http://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf (2011)

14. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. Management Information Systems Quarterly 26, 3 (2002)
15. Reifel, J., Hales, M., Pei, A., Blanter, A., Lala, S. and Bharadwaj, N.: The Internet of Things: Opportunities for Insurers, https://www.atkearney.com/digital-business/ideas-insights/featured-article/-/asset_publisher/Su8nWSQlHtbB/content/internet-of-things-opportunity-for-insurers/10192 (2014) (Accessed: 10.08.2016)
16. Reker, J.: Digitalisierung im Mittelstand, http://www2.deloitte.com/content/dam/Deloitte/de/Documents/Mittelstand/Digitalisierung-im-Mittelstand.pdf (2013) (Accessed: 12.08.2016)
17. Büst, R., Hille, M. and Schestakow, J.: Digital Business Readiness - Wie deutsche Unternehmen die Digitale Transformation angehen, https://www.crisp-research.com/report/digital-business-readiness-wie-deutsche-unternehmen-die-digitale-transformation-angehen/ (2015) (Accessed: 02.08.2016)
18. Fitzgerald, M., Kruschwitz, N., Bonnet, D., Welch, M.: Embracing Digital Technology: A New Imperative. Findings From the 2013 Digital Transformation Global Executive Study and Research Project by MIT Sloan Management Review & Capgemini Consulting. MIT Sloan Management Review, 1–12 (2013)
19. Westerman, G., Calmejane, C., Bonnet, D., Ferraris, P. and McAfee, A.: Digital Transformation: A Roadmap for Billion Dollar Companies, https://www.capgemini.com/resources/digital-transformation-a-roadmap-for-billiondollar-organizations (2011) (Accessed: 10.07.2016)
20. Guest, M.: Building your digital DNA - Lessons from digital leaders, https://www2.deloitte.com/content/dam/Deloitte/br/Documents/technology/deloitte-uk-building-your-digital-dna.pdf (2014) (Accessed: 02.08.2016)
21. Bloching, B., Leutiger, P., Oltmanns, T., Rossbach, C., Schlick, T., Remane, G., Quick, P. and Shafranyuk, O.: The digital transformation of industry, http://bdi.eu/media/user_upload/Digital_Transformation.pdf (2015) (Accessed: 30.10.2016)
22. Riemensperger, F., Hagemeier, W., Pfannes, P., Wahrendorff, M. and Feldmann, M.: Mut, anders zu denken: Digitalisierungsstrategien der deutschen Top500, https://www .accenture.com/de-de/service-deutschlands-top-500.aspx (2015) (Accessed: 12.08.2016)
23. Keen, P., Williams, R.: Value Architectures for Digital Business: Beyond the Business Model. Management Information Systems Quarterly 37, 642–647 (2013)
24. Markus, M.L., Loebbecke, C.: Commoditized Digital Processes and Business Community Platforms: New Opportunities and Challenges for Digital Business Strategies. Management Information Systems Quarterly 37, 649–654 (2013)
25. Pagani, M.: Digital Business Strategy and Value Creation: Framing the Dynamic Cycle of Control Points. Management Information Systems Quarterly 37, 617–632 (2013)
26. Grover, V., Kohli, R.: Revealing your Hand: Caveats in Implementing Digital Business Strategy. Management Information Systems Quarterly 37 (2013)
27. Oestreicher-Singer, G., Zalmanson, L.: Content or Community? A Digital Business Strategy for Content Providers in the Social Age. Management Information Systems Quarterly 37, 591–616 (2013)
28. Yoo, Y., Henfridsson, O., Lyytinen, K.: The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research. Information Systems Research 21, 724–735 (2010)
29. Barrett, M., Davidson, E., Fayard, A.-L., Vargo, S., Yoo, Y.: Being Innovative about Service Innovation: Service, Design and Digitalization. In: Proceedings of the 2012 International Conference on Information Systems (ICIS 2012), Panel Statement (2012)

30. Lusch, R.F., Nambisan, S.: Service Innovation: A Service-Dominant Logic Perspective. Management Information Systems Quarterly 39, 155–175 (2015)
31. Hylving, L., Henfridsson, O., Selander, L.: The Role of Dominant Design in a Product Developing Firm's Digital Innovation. Journal of Information Technology Theory and Application 13, 5 (2012)
32. Hylving, L., Schultze, U.: Evolving the Modular Layered Architecture in Digital Innovation: The Case of the Car's Instrument Cluster. In: Proceedings of the 2013 International Conference on Information Systems (ICIS 2013) (2013)
33. Chowdhury, S., Akesson, M.: A Proposed Conceptual Framework For Identifying The Logic Of Digital Services. In: Proceedings of the 15th Pacific Asia Conference on Information Systems (PACIS 2011), Paper 47 (2011)
34. Matt, C., Hess, T., Benlian, A.: Digital Transformation Strategies. Business & Information Systems Engineering 57, 339–343 (2015)
35. Gaskin, J., Thummadi, V., Lyytinen, K., Yoo, Y.: Digital Technology and the Variation in Design Routines: A Sequence Analysis of Four Design Processes. In: Proceedings of the 2011 International Conference on Information Systems (ICIS 2011), Paper 14 (2011)
36. Kuehne, K., Kosch, L., Cuylen, A.: Will XML-based Electronic Invoice Standards Succeed?-An Explorative Study. In: Proceedings of the 2015 European Conference on Informaion Systems (ECIS 2015), Paper 113 (2015)
37. Lu, B., Guo, X., Luo, N., Chen, G.: Corporate Blogging and Job Performance: Effects of Work-related and Nonwork-related Participation. Journal of Management Information Systems 32, 285–314 (2015)
38. Ceccagnoli, M., Forman, C., Huang, P., Wu, D.J.: Digital Platforms: When is Participation Valuable? Communications of the ACM 57, 38–39 (2014)
39. Dahl, A., Lawrence, J., Pierce, J.: Building an Innovation Community. Research-Technology Management 54, 19–27 (2011)
40. Hanelt, A., Piccinini, E., Gregory, R.W., Hildebrandt, B., Kolbe, L.M.: Digital Transformation of Primarily Physical Industries-Exploring the Impact of Digital Trends on Business Models of Automobile Manufacturers. In: Proceedings of the 12. International Conference on Wirtschaftsinformatik, pp. 1313–1327 (2015)
41. Mädche, A.: Interview with Wolfgang Gaertner on ‘‘Digitalization in Retail Banking: Differentiation and Standardization Through IT’’. Business & Information Systems Engineering 57, 83–85 (2015)
42. Spann, M.: Interview with Jörg Lübcke on "Digitalization of Business Models in the Media Industry". Business & Information Systems Engineering 5, 199–201 (2013)
43. Winter, R.: Modelle, Techniken und Werkzeuge im Business Engineering. In: Österle, H., Winter, R. (eds.) Business Engineering. Auf dem Weg zum Unternehmen des Informationszeitalters, pp. 87–118. Springer, Oldenbourg, Munich (2003)
44. Aier, I.S., Winter, R.: Virtual decoupling for IT/business alignment–conceptual foundations, architecture design and implementation example. Business & Information Systems Engineering 1, 150–163 (2009)
45. Buhl, H.U., Kaiser, M.: Herausforderungen und Gestaltungschancen aufgrund von MiFID und EU-Vermittlerrichtlinie. Zeitschrift für Bankrecht und Bankwirtschaft 20, 43–51 (2008)
46. Huber, J.: E-Commerce: Megatrend „Social, Local, Mobile". In: Knoll, M., Meinhardt, S. (eds.) Mobile Computing: Grundlagen – Prozesse und Plattformen – Branchen und Anwendungsszenarien, pp. 65–79. Springer Fachmedien Wiesbaden, Wiesbaden (2016)
47. Kleindienst, D.: Social Media Analytics: Wie die Ausrichtung an den Unternehmenszielen gelingt. HMD Praxis der Wirtschaftsinformatik 53, 736–747 (2016)

48. Frank, U.: Multi-Perspective Enterprise Modeling (Memo) Conceptual Framework and Modeling Languages. In: Proceedings of the 35th Annual Hawaii International Conference on System Sciences, pp. 1258–1267 (2002)
49. Winter, R., Fischer, R.: Essential Layers, Artifacts, and Dependencies of Enterprise Architecture. Journal of Enterprise Architecture (2007)
50. Neumeier, A., Wolf, T. and Oesterle, S.: The Manifold Fruits of Digitalization - Total Reference List, https://github.com/digitalizationpaper/WI17/blob/master/References.pdf (2016)
51. Davenport, T.H.: Process innovation. Reengineering work through information technology. Harvard Business School Press, Boston, Mass. (1993)
52. Röglinger, M., Pöppelbuß, J., Becker, J.: Maturity Models in Business Process Management. Business Process Management Journal 18, 328–346 (2012)
53. Kano, N., Seraku, N., Takahashi, F., Tsuji, S.: Attractive Quality and must-be Quality. Journal of the Japanese Society for Quality Control 14, 39–48 (1984)
54. Mette, P., Moser, F., Fridgen, G.: A Quantitative Model for Using Open Innovation in Mobile Service Development. In: Proceedings of the 11th International Conference on Wirtschaftsinformatik, 1, pp. 71–86 (2013)
55. Fornell, C., Johnson, M.D., Anderson, E.W., Cha, J., Bryant, B.E.: The American Customer Satisfaction Index: Nature, Purpose, and Findings. Journal of Marketing 60, 7–18 (1996)

Rebound Effects in Cloud Computing: Towards a Conceptual Framework

Michael Adelmeyer[1], Marc Walterbusch[1], Peter Biermanski[1], Kai Seifert[1], and Frank Teuteberg[1]

[1] University of Osnabrück, Institute of Accounting and Information Systems, Osnabrück, Germany
{michael.adelmeyer,marc.walterbusch,pbiermanski,kaseifert, frank.teuteberg}@uni-osnabrueck.de

Abstract. Rebound effects have been discussed in various disciplines. In the information and communication technology sector, this topic is still insufficiently studied. Basically, a rebound effect is a feedback mechanism, as a result of which savings from efficiency improvements are not or only partially realized. Due to the potential of cloud computing for efficiency improvements, not only in terms of energy efficiency, but also in terms of organizational resources in general, we describe rebound effects in this context by means of a systematic literature review and a case study. Our results provide a framework to categorize and identify potential rebound effects in cloud computing. The understanding of rebound effects and their influence on the various organizational resources (e.g., server hardware, human resources or IT know-how), is important for managers to sustainably decide for or against the adoption, integration and roll out of cloud computing services.

Keywords: Rebound Effects, Cloud Computing, Literature Analysis, Case Study, Conceptual Framework.

1 Introduction

In recent years, the cloud computing (CC) technology has emerged as a new computing paradigm and has gained increasing attention due to its remarkable advantages, e.g., reducing costs and complexity along with increasing flexibility [1, 2]. In its core, CC resorts to existing technologies like grid computing, virtualization and web services for the online delivery of scalable IT services, commonly on the basis of a pay-per-use pricing model [3]. However, when strategically outsourcing a supportive or even a core process into the cloud (so called “cloudsourcing”), it is mandatory that the management of the organization has a clear understanding of all influencing as well as influenced factors. One of these factors is the rebound effect, which diffuses alongside with the efficiency improvements of CC and which describes that expected savings resulting from these improvements cannot be realized, due to a growth in consumption of the underlying or different resources.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Adelmeyer, M.; Walterbusch, M.; Biermanski, P.; Seifert, K.; Teuteberg, F. (2017): Rebound Effects in Cloud Computing: Towards a Conceptual Framework, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 499-513

Rebound effects have been discussed in various disciplines, e.g., psychology, but in the information and communication technology (ICT) sector this topic is still insufficiently investigated [4, 5]. Moore's law constitutes a commonly referenced example of rebound effects in the ICT field. It predicts that the microchip performance per cost unit doubles every 18 months, which makes ICT hardware widely affordable, but in turn leads to an increase of consumption [5, 6]. Existing literature on rebound effects in the context of CC mainly focuses on an energy saving perspective [7–9]. For example, Walnum and Andrae (2016) estimate that rebound effects possibly offset energy savings at the macro-level due to the increased amount of data transferred [8]. In the field of ICT, potential savings do not only result from energy efficiency improvements, but also from organizational resources, such as workforce [4]. Therefore, we expand the focus from a narrow energy saving perspective to efficiency improvements of general organizational resources in the context of CC. For example, if an organization outsources servers and corresponding storage into the cloud, the decreased costs as well as the enhanced provisioning of storage may lead employees to make use of the new storage facilities more extensively than before. This in turn would exhaust the expected savings, which consequently remain unrealized.

To avoid possible rebound effects that would absorb the resource-saving impact of CC, a basic understanding must be established [4]. With this in view, we conducted a systematic literature review focusing on the definition of rebound effects. Subsequently, we describe the phenomenon of rebound effects in CC by means of a case study, in which we accompany an organization during a typical outsourcing process into the cloud. Based on the findings, we provide a framework consisting of a definition of rebound effects in general (REGE) and rebound effects in CC (RECC), a conceptual model and a morphological analysis of RECC to identify, categorize and understand potential rebound effects in cloud computing. In practice, the final decision on the usage of CC services has a significant influence on the various organizational resources (e.g., server hardware, human resources or IT know-how). The understanding of surrounding factors such as the rebound effects is vitally important for decision makers to sustainably decide for or against the adoption, integration and roll out of CC services. From the scientific perspective, the framework constitutes a starting point for further investigation.

The corresponding research questions (RQs) we seek to answer are the following:

RQ1: How can rebound effects in cloud computing be conceptualized?

RQ2: Which organizational resources are potentially affected by rebound effects in cloudsourcing scenarios?

Before we answer these RQs, we outline our methodical approach and provide theoretical background on rebound effects in general. Second, we present a definition, a conceptual model and a morphological analysis concerning rebound effects in CC on the basis of a literature analysis and a case study approach. Finally, we discuss our research findings, line out implications for theory and practice and conclude with limitations as well as future research implications.

2 Research Methods

2.1 Systematic Literature Review

In order to identify and analyze relevant research on rebound effects, we carried out a systematic literature review according to vom Brocke et al. (2009) [10]. The systematic literature review provides the basis for the definition of REGE (cf. section 3) and consequently indirectly for the conceptual modelling and the morphological analysis of RECC (cf. section 4.1). In order to achieve high-quality results, we limited the search to the top 25 (out of 109) journals included in the AIS (2014) ranking, all A- and B-ranked journals and conferences according to the VHB (2008) ranking, as well as all A-ranked journals and conferences included in the WKWI (2008) ranking [11–13]. All the above meta-rankings of international information systems journals and conferences are widely accepted. Additionally, due to their diversified programs, we also considered the well-known conferences Americas Conference on Information Systems (AMCIS) and the Hawaii International Conference on System Sciences (HICSS) to be relevant.

For the queries, we defined the following search terms and phrases (rebound effect*, rebound-effect*, reboundeffect*, rebound mechanism*, rebound-mechanism*, reboundmechanism*, boomerang*). The wildcards used ensure the identification of related terms (e.g., ‘rebound effect’ or ‘rebound effects’ for ‘rebound effect*’). In this search, we deliberately left out the refining search term cloud computing, as the achieved results will naturally already contain the contributions in this field. With the listed search terms we searched the titles and abstracts of the previously mentioned publication organs and found only one paper. Since this initial search focused on IS publication organs, the implication by Gossart (2014) that the phenomenon of rebound effects is still largely unexplored in the field of ICT is verified [4]. Due to this result, we extended the initial search to the databases EBSCOhost, SpringerLink and Google Scholar, which led to a total of 2.325 results (EBSCOhost: 293, Google Scholar: 572, SpringerLink: 1.460). As these results are sorted by relevance in the corresponding databases, we focused on the first 100 results for each query in each database. After reading the titles and abstracts, we identified 54 relevant results (EBSCOhost: 21, Google Scholar: 16, SpringerLink: 17). A forward and/or backward search based on these results was not conducted since the initial literature search already distinctly focused on IS literature.

2.2 Case Study

The case study research method aims to study a single, contemporary phenomenon (e.g., an application, a technology or a decision) within its real-life context (e.g., in an organization) by investigating multiple sources of evidence (e.g., archival records and physical artifacts) over a logical time frame [14–19]. All these characteristics are especially important when the boundaries between phenomenon and context are not clearly evident [16, 19]. Case studies facilitate the understanding of the nature and complexity of the processes taking place [14, 17, 18]. Besides, they are particularly suitable for the investigation of information systems development, implementation and use within organizations [18, 20, 21].

As case studies can be distinguished by the research objective (descriptive, explorative, explanatory), research epistemology (positivist, interpretive, critical), research method (qualitative, qualitative/quantitative), case design (single, multiple) and unit of analysis (holistic, embedded) [20, 22], a variety of approaches exist to carry out case studies [15, 17, 18]. Our study generally follows the synthesized case research procedures by Dubé and Paré (2003), Gable (1994), Newell and Simon (1972) and Recker (2013), as depicted in Figure 1 [19, 22–24]. Instead of recommending specific case study methods, these procedures define fundamental elements and sort them in order to deduce theoretical statements [25].

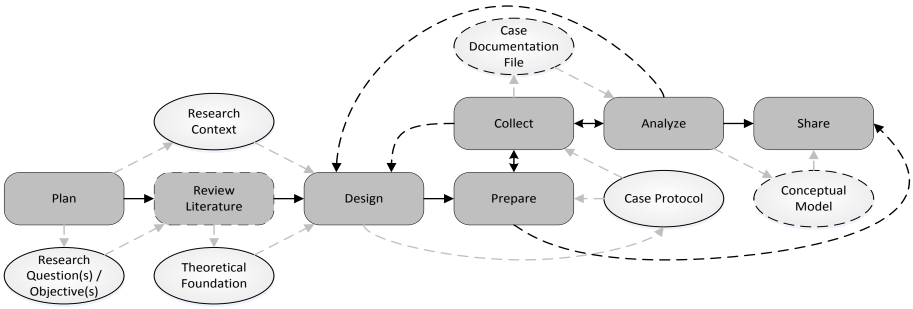


Figure 1. Synthesized case research procedure in accordance with Dubé and Paré (2003), Gable (1994), Newell and Simon (1972) and Recker (2013) [19, 22–24]

The single case study approach is particularly suitable for the investigation of previously unchallenged phenomena or issues at the beginning of theory generation. This is in line with our research objective to describe the emerging phenomenon of RECC (cf. section 1 & 2.1). A multiple case set-up, however, allows the researcher to test a theory [21, 22]. Therefore, this approach is appropriate when a theory has to be validated and a framework has to be evaluated. We try to explore the object of interest within its real-life context in an organization without defining any (in)dependent variables a priori [23, 26].

To be precise, we explore RECC at NOZ Medien, a media company in Northern Germany. NOZ Medien has more than 950 employees who are spread across nine locations. It has a daily readership of over half a million people. We have chosen NOZ Medien for our case study, since the decision makers confirmed their general interest in outsourcing processes into the cloud. Furthermore, top-level management assured full support but had not yet taken any steps towards "cloudsourcing". During the period from October 2013 until September 2014, we accompanied IT decision makers at NOZ Medien during a typical "cloudsourcing" process, which was predetermined by the company, from the (i) idea to outsource, (ii) as-is analysis, (iii) requirements definition, (iv) market analysis and (v) simulation to the point of (vi) transition or abortion. Since during the step 'transition or abortion' the IT decision makers at NOZ Medien came to the conclusion not to outsource the considered system into the cloud, the steps (vii) contract negotiations, (viii) migration, (ix) operation, (x) evaluation and (xi) optimization or discarding have not been run through. Given the fact that we factored in the possibility of cancellation of the outsourcing process when designing the case study and that we follow a conceptual approach for the definition of RECC, the impact on the conceptual framework is limited. In the present case study, the idea to outsource targets the email system at NOZ Medien. This system constitutes an autonomous unit inside the IT system landscape and is based on Microsoft Exchange 2007.

During the project steps conducted, we made use of several data collection methods in order to capture the contextual complexity and to obtain a substantial set of data [18]: (i) documentation (e.g., infrastructure plan), (ii) archival records (e.g., organization chart), (iii) interviews (e.g., with the IT decision makers) and (iv) physical artefacts (e.g., IT landscape); (v) direct observations were not included. These data are analyzed in a qualitative manner following an interpretive approach according to Myers (1997), who states that "interpretive studies generally attempt to understand phenomena through the meanings that people assign to them […]" and are "[…] aimed at producing an understanding of the context of the information system, and the process whereby the information system influences and is influenced by the context" [20, 27].

3 Rebound Effects in General

There is no uniform definition of the term rebound effect, it is rather discussed controversially [28, 29]. The primal characterization of this effect can basically be ascribed to Jevons (1865), who suggests that the efficiency improvement of coal-fired steam engines leads to a higher instead of a lower consumption of coal [30]. The reason for this is the reduction of the effective cost of coal realized through the higher efficiency [30, 31]. Sorrell et al. (2009) further distinguish between direct and indirect rebound effects, whereas economy-wide rebound effects reflect the sum of the two aforementioned [32]. The term rebound effect is not only used and discussed in disciplines related to sustainability but also in several others [32, 33], for example, in biology, physics, chemistry and psychology as well as in economic contexts, e.g., logistics. This highlights the variety of application areas on the one hand, and the multitude of existing rebound effects on the other. Gossart (2014), Plepys (2002) and Hilty et al. (2006) draw attention on the effects in the ICT sector [4, 6, 34]. They examine the existence of rebound effects in information and communication technology and their impact on the economy, whereas this topic still remains insufficiently explored in this area. Andrae (2013), Sedlacko et al. (2014) and Walnum and Andrae (2016) explore rebound effects in the context of energy savings in cloud computing [7–9].

Mathematically, rebound effects can most simply be described as the quotient of the difference between expected and realized savings (numerator), which is subject to an (unintended) alteration of the consumption of the resource regarded or different resources, and the expected savings (divisor) [35–37]:

$$Rebound\ effect\ (\%) = \frac{100*(expected\ savings - realized\ savings)}{expected\ savings}$$

Figure 2. Mathematical calculation of a rebound effect

In case the realized savings are smaller than the expected savings, this results in a positive rebound effect (percentage). If no savings can be realized, the efficiency improvement even results in a surplus of consumption (backfire); the rebound effect would then be greater than 100% (full exhaustion). Mathematically, in case more savings are realized than expected, also a negative percentage is possible.

Based on the systematic literature review, we build a concept matrix according to Webster & Watson (2002) in order to summarize how rebound effects are characterized [cf. http://bit.ly/RECCWI17] [38]. Out of the initial 54 results, 47 papers contained concepts characterizing rebound effects in several ways. In order to uncover and name the underlying concepts, we applied a conceptual content analysis approach when analyzing the literature [22], placing a strong focus on the general definition of rebound effects. In total, the following eight characterizing concepts of rebound effects in general were extracted from the identified literature: efficiency improvement; growth in consumption; direct/indirect; micro/macro level; offset; unintended; short-/long term; behavioral response.

The main cause for rebound effects is an efficiency improvement, which leads to a behavioral response, namely a (unintended) growth in consumption and an offset of the initial savings on a short- or a long term [31]. The scope of rebound effects can be distinguished between micro- (single actor) as well as macro-level (interaction of several actors) and further between direct (investment in the same resource or service) as well as indirect (investment in alternative/substituting resource(s) or service(s)) [31–33]. In summary,

A rebound effect describes the unrealized [(over)exhausted] saving of a resource in consequence of a (unintentional) growth in consumption, whereas the saving of this resource could have been expected and would have been possible based on the efficiency improvement concerning the use of this resource. A distinction is made between direct and indirect rebound effects, which affect the micro- as well as macro-level in the short- or long-term.

In the context of rebound effects, the commonly considered resource is energy in the form of, e.g., electricity or gasoline. In the following, the term resources will implicitly include any other forms of resources. Since energy is indirectly affected by the (efficient) use of time, time itself and as consequence workforce can also be seen as a resource [4, 5]. In addition to the main cause for rebound effects, namely an efficiency improvement, numerous secondary causes (also called “effects”) are discussed in literature, such as financial (caused by cost savings, e.g., income, reinvestment and market price effect), material (caused by the energy that is used for the production and provision of goods and services, e.g., embodied energy, consumption accumulation and new markets effect) or psychological effects (caused by increased demand as a result of moral justification, e.g., moral hazard, moral leaking and moral licensing effect) [31].

4 Rebound Effects in Cloud Computing

4.1 Definition and Conceptualization

The efficiency increase described in Moore's law concerning the increase of computational performance per cost unit leads in turn to a broad diffusion and a higher demand of the technology and consequently to an increase in energy consumption by ICT hardware [5, 6]. Due to its characteristics like resource pooling and virtualization, the cloud computing technology has the inherent potential for efficiency gains of resources, such as energy [8, 39].

Therefore, in a first step, the definition of rebound effects is transferred to CC. Starting point is the definition of REGE, which was derived based on the concept matrix resulting from the literature analysis (cf. section 3). Equally important is the de-facto definition for CC by Mell and Grance (2011), who define CC as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" [40]. Furthermore, cloud computing can be distinguished between three service models (Software, Platform and Infrastructure as a Service) and different deployment models, such as private, community, public or hybrid clouds [40]. The results of the case study constitute another starting point for the definition of RECC. The cause for a REGE is an efficiency improvement. CC can lead to such efficiency improvements in various respects. In the following, one example for each of the five essential characteristics of CC [40] will be given: (i) on-demand self-service leads to time savings in the provision of computing capabilities; (ii) broad network access leads to independencies for using CC services via heterogeneous client platforms; (iii) resource pooling leads to savings in terms of maintenance, replacement and repair; (iv) rapid elasticity leads to a constant, appropriate performance; (v) measured service leads to a full control of costs.

The resource underlying a rebound effect can be of various natures: concerning REGE, the focus is generally put on the resource energy. In CC, the efficiency increase results from the use of cloud computing services, which in turn affects resources that are specific to the organization, e.g., liquid funds, working time and corresponding assets like hardware. Moreover, when evaluating a RECC, the overall corporate objectives (organizational perspective) have to be considered. Assuming that an organization outsources its servers into the cloud using Infrastructure as a Service (IaaS) in the form of virtualized resources from a CC provider, the servers used so far as well as the human resources, who administrated these servers (e.g., technicians and to some extent the former administrators), are no longer required. The residual value of the hardware and the monetized released working time, after deducting the costs for the CC service, can be interpreted as savings. Depending on the corporate objectives, these savings can or better should be used continuously. If the objectives comprise the steady growth of human resources [41], the investment of the residual value of the hardware as well as the transformation of the released working time into a new position, will mathematically lead to a positive rebound effect (exhaustion of the savings). However, only if these reinvestments are directly caused by or related to the adoption of CC services and diminish the realized savings in comparison to the expected savings, the result is a RECC. In case the investments are made based on the realized savings generated through the efficiency improvement of the adoption of CC, e.g., in order to increase production by means of new employees and production machinery, the savings were already realized and therefore do not result in a RECC. As long as the consumption of or investment in the same or (an) alternative resource(s) is in line with the corporate objectives, the rebound effect can be seen as positive from an organizational perspective. The situation would be different, however, if the corporate objectives comprised the gradual build-up of capital reserves and a high degree of automation. In this case, the said investments would mathematically still lead to a positive rebound effect, but as not in line with the corporate objectives, it would have to be assessed negatively from an organizational perspective.

Consequently, a RECC is depended on the context of an organization's intention or expectation. This interplay between the intentions or the expectations of an organization and the actual effect can be seen in Table 1.

Table 1. Rebound effects in cloud computing in the context of intention or expectation and effect in accordance with Hilty (2008) [41]

<table>
<tr><td></td><td></td><td colspan="2">intention or expectation</td></tr>
<tr><td></td><td></td><td>reduce input</td><td>increase output</td></tr>
<tr><td rowspan="3">effect</td><td rowspan="2">input does not fall</td><td>rebound effect in the narrow sense</td><td>intended growth</td></tr>
<tr><td colspan="2">rebound effect in the broad sense</td></tr>
<tr><td>output does not rise</td><td>saving of input</td><td>unintended stagnation</td></tr>
</table>

Depending on the point of view of an economic actor (private households, organizations, state), the intentions or expectations that cause rebound effects differ. Since an organization is generally interested in growth (increase in output), rebound effects in the narrow sense occur if an organization has the intention or expectation to accumulate savings (reduce input). Rebound effects in the broad sense arise under the intention of growth, leading to a disproportional resource consumption [41].

In addition, RECC can be distinguished between direct and indirect effects. For example, savings resulting from outsourcing a process into the cloud can directly be reinvested in the underlying resource saved or indirectly, e.g., in additional workforce managing the outsourced service. Furthermore, the effect level can be distinguished between micro (within the organization) and macro (outside the organization).

Based on the definition of REGE (cf. section 3) and the previous discussion, the following definition for RECC is derived:

A rebound effect in cloud computing describes the unrealized [(over)exhausted] saving of an organizational resource in consequence of a (unintentional) growth in consumption, whereas the savings could have been expected and would have been possible based on the efficiency improvement resulting from the use of cloud *computing services. The organizational assessment of the effect is depending on the corporate objectives. A distinction is made between direct and indirect cloud computing rebound effects, which affect the micro- as well as macro-level in the short- or long-term.*

Concerning RECC, the organizational context as well as internal decisions are focused, which are basically influenced by the corresponding corporate objectives. The conceptual model depicted in Figure 3, which is based on the literature analysis and the results of the case study, summarizes the components of a RECC.

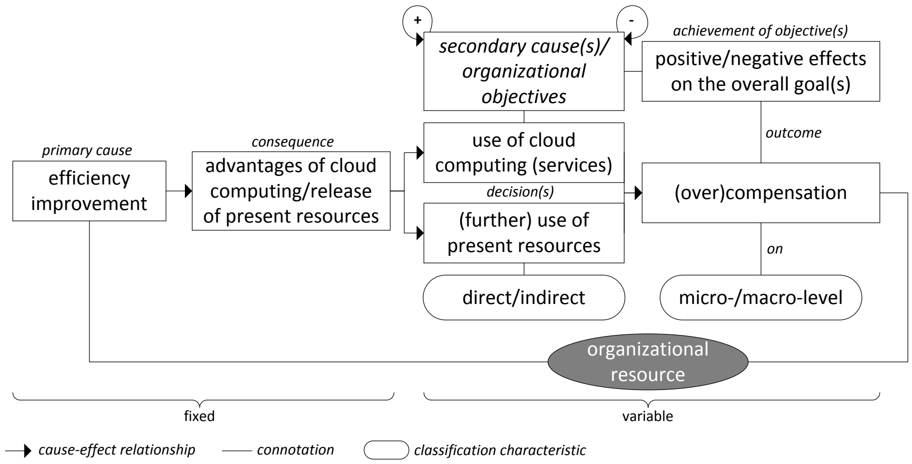


Figure 3. Conceptual model of a rebound effect in cloud computing

Finally, a morphological analysis has been conducted which combines the results of the concept matrix resulting from the literature analysis and the case study in order to investigate the “totality of relationships contained in multi-dimensional, usually non-quantifiable problem complexes”, in this case RECC [42] (cf. Table 2). The morphological analysis allows for the categorization of potential RECC, whereas the conceptual model primarily depicts connotations and causalities.

Table 2. Morphological analysis of rebound effects in cloud computing

<table>
<tr><td>cause and effect</td><td colspan="4">Saving (of an organizational resource) due to the use of a cloud computing service (IaaS, PaaS, SaaS), leading to an unintended effect on (other) organizational resources (growth in consumption)</td></tr>
<tr><td>(further) use of resources</td><td colspan="2">direct</td><td colspan="2">indirect</td></tr>
<tr><td>effect level</td><td colspan="2">micro</td><td colspan="2">macro</td></tr>
<tr><td>compensation</td><td>negative</td><td colspan="2">partially</td><td>backfire</td></tr>
<tr><td>period</td><td colspan="2">short-term</td><td colspan="2">long-term</td></tr>
<tr><td>assessment</td><td colspan="2">mathematical</td><td colspan="2">organizational (corporate objectives)</td></tr>
</table>

In summary, RECC are unrealized savings of an organizational resource caused by efficiency improvements, which arise from the migration to CC services as well as from the (unintended) effects on (other) organizational resources. Since the case study targets a SaaS solution and no distinction in existing literature on RECC could be identified, no further differentiation of RECC concerning service and deployment models could be made. Regarding the assessment of a RECC, a mathematical and an organizational viewpoint need to be distinguished. For example, it may be the case that a mathematically calculated partial compensation, thus a partial rebound effect, does not, only partially or disproportionately affect the corporate objectives. However, a mathematically calculated compensation can be negative (realization of more savings than expected), which can in turn lead to a disproportionate effect on the organizational goals, e.g., if the overall goal of an organization was a complete reinvestment or growth or to retain jobs.

4.2 Identification of Affected Resources

Since a REGE describes the efficiency improvement with respect to resources such as energy or the saving of these, organizational resources have been derived for the discussion of RECC. Basic IT-related resources, which are typically used in an organization, have been identified by means of a classification of accounts for industrial enterprises [43]. Due to the fact that this classification is directed towards the documentation and structuring of organizations' assets, it is well suited for our purpose. We concentrated on those basic resources that can directly or indirectly be influenced by changes in IT. For example, the outsourcing of servers might have a definable influence on the expenses for (external) maintenance; however, there are other effects, e.g., on expenses for advertising, that are difficult to assess. In order to expand the identified basic resources by special resources, that are also affected by CC, the three CC service models – Infrastructure (IaaS), Platform (PaaS) and Software as a Service (SaaS) – have been investigated [40]. Additionally, we held workshops with IT-management and conducted two expert interviews in May 2014 during the case study to identify potentially affected resources. Furthermore, the transcripts and the company's documentation (e.g., IT landscape, infrastructure plans, balance sheets) were analyzed qualitatively via inductive category development. Finally, all results have been aggregated (cf. Table 3).

Table 3. Organizational resources that are potentially (directly or indirectly) affected by changes in IT (in the short- or long-term)

Aggregated Category	**Organizational Resources**
IT related	employees, IT know-how, space, server software, energy, repair material, internal data and information, maintenance
Hardware & Systems	backup and recovery, server hardware, adjacent assistive systems, infrastructure and connection, air conditioning, devices and clients
General	money, customers and reputation, suppliers and partners, non-IT know-how

The categories and underlying organizational resources depicted in Table 3 are directly or indirectly affected by changes in IT, e.g., resulting from a possible cloudsourcing decision, and are therefore subject to potential RECC. For example, one expert states that migrating to the cloud would "save maintenance workload and costs". The major resource that is expected by the experts to be directly affected is money ("outsourcing to the cloud is way cheaper in comparison to an in-house solution"). If the expectation or intention regarding an organizational resource, in this case, to save money or maintenance, cannot or only partially be realized by an adoption of cloud services, this would lead to a rebound effect. Furthermore, depending on the organizational goals, resource saving potentials might be evaluated negatively from an organizational point of view. For example, the saving potential of workforce or rather employees contradicts the overall social goal of NOZ Medien to provide stability for employees and to safeguard jobs. Therefore, the experts at NOZ Medien explain that affected employees would be entrusted with other responsibilities in order to retain jobs and to prevent the loss of valuable IT know-how.

On the foundation of the organizational resources presented in Table 3 and specifically for the present case study at NOZ Medien, expected effects on organizational resources and goals resulting from a cloudsourcing decision were identified. In the case study, the existing email system is to be replaced by a user- or better to say mailbox-based service from a CC provider (SaaS). Therefore, the expected effects have only been derived and evaluated a priori. First, the corporate objectives at NOZ Medien were identified in an interview with the management and mapped against the organizational resources (e.g., the economic factor "the highest profit possible" is connected to the general organizational resource "money"). Second, expected effects on organizational resources were identified and evaluated together with the experts in workshops. The resulting analysis can be found in Table 4. As a result of the case study, effects have to be assessed depending on the organizational goal, as elucidated by the resource "employees".

Table 4. Expected effects of cloudsourcing on organizational resources at NOZ Medien

organizational goals	organizational resource (goal/expected effect)
"highest profit possible"	• money (+/+)
"continued growth of the company" & "provide a perspective and future for Employees" & "provide stability for employees and safeguard jobs"	• employees (+/-)
"efficiency and sensible use of resources"	• hardware & systems (-/-) • server software (-/-) • employees (-/-) • space (+/+) • energy (-/-) • maintenance (-/-) • money (+/+) • IT know-how (o/-)

5 Discussion and Conclusion

Potential RECC can arise in any area of an organization in which outsourcing into the cloud results in resource saving potentials. In contrast to the traditional view on rebound effects, which generally focuses on a narrow energy saving perspective [7–9], we expanded the focus and added an organizational perspective [41], in which the adoption process of a cloud service takes place. In this context, the intentions or expectations of organizations can lead to different assessments of rebound effects from a mathematical (compensation of the savings) in comparison to an organizational (compliance with the corporate objectives) perspective. Furthermore, rebound effects in CC do not necessarily arise from energy savings. Rather, the effects can be the results of savings of any organizational resource, which is subject to expected efficiency improvements in a cloudsourcing scenario (cf. RQ2).

The framework consisting of a REGE/RECC definition, a conceptual model, and a morphological analysis of RECC provides a foundation to identify, categorize and understand potential rebound effects in cloud computing (cf. RQ1). Furthermore, our results confirm the complex dependencies that are not always identifiable at first glance, which can involve many possible causes and manifestations of RECC. Consequently, there is an urgent need to research the existence and impact of RECC. In line with Gossart (2014), we argue that future research should focus on rebound effects in the ICT context in order to understand and mitigate the underlying causalities [4]. This broadened understanding can be very helpful when exploring the decision-making process as well as the involved influencing and influenced factors. Consequently, future research should also comprise the investigation of RECC in the short- as well as long-term. In order to accomplish this requirement, we provided a fundamental framework. For practice, the awareness of RECC can facilitate to (at least approximately) assess the impact of rebound effects even prior to outsourcing, since various organizational resources are significantly affected by a cloudsourcing decision. The awareness of factors surrounding such a decision is vital for making a sustainable judgement.

As any research endeavor, our findings also need to be viewed in the light of some limitations. First and foremost, the decision process for resource allocation underlying a rebound effect in an organization differs – at least to some extent – from other research areas of rebound effects cited. Unlike what is the case in, e.g., natural sciences, the decisions in organizations cannot always be fully predicted and do not necessarily follow general principles. Organizational decisions are mostly based on individual or community decisions and are limited by given budgets [6]. To consider this, these organizational goals were taken into account when developing the framework in accordance to Hilty (2008), who further differentiates between rebound effect perspectives of private households, enterprises and states [41].

The overall target of the paper is to present the results obtained from the systematic literature review as well as the case study.[1] Although case studies are considered to be an appropriate method to capture the richness of organizational contexts – in this case

[1] The case documentation file (cf. Figure 1) will be provided to interested readers.

of a German media company – the conclusions drawn cannot be unreservedly generalized [23]. Neither do we assert that the understanding of RECC would substitute a decision-making process for any given "cloudsourcing" project. Instead, the intention of our research is to arouse awareness that unintended rebound effects potentially exist in CC, which can affect organizational resources. By generalizing our results in the forms of definitions, a conceptual model as well as a morphological analysis of RECC, we already mitigate this limitation to some extent. Moreover, since the focus of the single case study was put on outsourcing an email system, further results may arise when investigating the outsourcing of other systems into the cloud. Therefore, rebound effects in different systems, application areas and organizations in the context of cloud computing or else in other ICT contexts are subject to future research. Especially regarding the various cloud service or deployment models, potential differences and unique characteristics concerning rebound effects need to be investigated.

Moreover, since NOZ Medien at the point of transition or abortion decided not to outsource the email system into the cloud, we have not been able to verify potential rebound effects by means of direct observation. Our study rather presents implications of RECC, however, they have only been derived a priori. Hence, our conceptual framework requires further validation and evaluation, for example via a multi case setting, to measure occurring RECC. Multiple case study setups allow for the testing of theories, whereas single case setups are preferable to identify and describe previously unchallenged or emerging phenomena, as executed in this study [22]. Furthermore, potential single case biases could be eliminated and multiple CC service models and the variety of factors influencing RECC could be investigated. To this point, our work provides the basis for future quantitative measurement of RECC. All in all, in order to improve the understanding of CC and all surrounding factors, we encourage researchers to extend the focus on technical advancements to take other influencing and influenced factors into account.

Acknowledgement

The authors thank all involved parties at NOZ Medien, especially Mr. Krum, Mr. Gödde, Mr. Gerdes and Mr. Granitza for their valuable insights, help and substantive support during the research process. Furthermore, the authors thank Mrs. Imhorst as well as the reviewers for their constructive feedback.

References

1. Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R.H., Konwinski, A., Lee, G., Patterson, D.A., Rabkin, A., Stoica, I., Zaharia, M.: A view of cloud computing. Communications of the ACM. 53, 50–58 (2010)
2. Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., Ghalsasi, A.: Cloud computing - The business perspective. Decision Support Systems. 51, 176–189 (2011)

3. Walterbusch, M., Martens, B., Teuteberg, F.: Exploring Trust in Cloud Computing: A Multi-Method Approach. In: Proceedings of the 21st European Conference on Information Systems (ECIS). Utrecht (2013)
4. Gossart, C.: Rebound Effects and ICT : A Review of the Literature. In: Hilty, L.M. and Aebischer, B. (eds.) ICT Innovations for Sustainability. Advances in Intelligent Systems and Computing. Springer International Publishing (2014)
5. Hilty, L.M., Lohmann, W., Huang, E.M.: Sustainability and ICT – An overview of the field. Notizie di Politeia. 27, 13–28 (2011)
6. Plepys, A.: The grey side of ICT. Environmental Impact Assessment Review. 22, 509–523 (2002)
7. Sedlacko, M., Martinuzzi, A., Dobernig, K.: A Systems Thinking View on Cloud Computing and Energy Consumption. Proceedings of the 2nd International Conference on ICT for Sustainability (ICT4S). 95–102, Stockholm (2014)
8. Walnum, H.J., Andrae, A.S.G.: The Internet: Explaining ICT Service Demand in Light of Cloud Computing Technologies. In: Santarius, T., Walnum, H.J., and Aall, C. (eds.) Rethinking Climate and Energy Policies: New Perspectives on the Rebound Phenomenon. pp. 227–241. Springer International Publishing, Switzerland (2016)
9. Andrae, A.S.G.: Comparative micro life cycle assessment of physical and virtual desktops in a cloud computing network with consequential, efficiency, and rebound considerations. Journal of Green Engineering. 3, 193–218 (2013)
10. vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the giant: on the importance of rigour in documenting the literature search process. In: Proceedings of the 17th European Conference on Information Systems (ECIS). Verona (2009)
11. AIS: MIS Journal Rankings (2014)
12. VHB: Teilranking Wirtschaftsinformatik und Informationsmanagement (2008)
13. WKWI: WI-Orientierungslisten. Wirtschaftsinformatik. 50, 155–163 (2008)
14. Palvia, P., Mao, E., Salam, A.F., Soliman, K.S.: Management Information Systems Research: What's There in a Methodology? Communications of the Association for Information Systems. 11, 289–309 (2003)
15. Bonoma, T.V.: Case Research in Marketing: Opportunities, Problems, and a Process. Journal of Marketing Research. 22, 199–208 (1985)
16. Yin, R.K.: Case Study Research: Design and Methods. Sage Publications, Thousand Oaks, CA (2014)
17. Kaplan, B., Duchon, D.: Combining Qualitative and Quantitative Methods in Information Systems Research: A Case Study. MIS Quarterly. 12, 571–586 (1988)
18. Benbasat, I., Goldstein, D.K., Mead, M.: The Case Research Strategy in Studies of Information Systems. MIS Quarterly. 11, 369–386 (1987)
19. Dubé, L., Paré, G.: Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations. MIS Quarterly. 27, 597–635 (2003)
20. Myers, M.D.: Qualitative Research in Information Systems. MIS Quarterly. 21, 241–242 (1997)
21. Darke, P., Shanks, G., Broadbent, M.: Successfully completing case study research: combining rigour, relevance and pragmatism. Information Systems Journal. 8, 273–289 (1998)
22. Recker, J.: Scientific Research in Information Systems: A Beginner's Guide. Springer, Heidelberg (2012)
23. Gable, G.G.: Integrating case study and survey research methods: an example in information systems. European Journal of Information Systems. 3, 112–126 (1994)

24. Newell, A., Simon, H.A.: Human Problem Solving. Prentice-Hall, Englewood Cliffs, NJ (1972)
25. Williams, C.K., Karahanna, E.: Casual Eplanation in the Coordinatin Process: A Critical Realist Case Study of Federated IT Governance Structures. MIS Quarterly. 37, 933–964 (2013)
26. Kaplan, B., Maxwell, J.A.: Qualitative Research Methods for Evaluating Computer Information Systems. In: Anderson, J.G., Aydin, C.E., and Jay, S.J. (eds.) Evaluating Health Care Information Systems: Methods and Applications. pp. 45–68. Sage, Thousand Oaks, CA (1994)
27. Walsham, G.: Interpreting Information Systems in Organizations. Wiley, Chichester (1993)
28. Barker, T., Dagoumas, A., Rubin, J.: The macroeconomic rebound effect and the world economy. Energy Efficiency. 2, 411–427 (2009)
29. Binswanger, M.: Technological progress and sustainable development: what about the rebound effect? Ecological Economics. 36, 119–132 (2001)
30. Jevons, W.S.: The Coal Question. Macmillan and Co., London (1865)
31. Santarius, T.: Green Growth Unraveled. How Rebound Effects Baffle Sustainability Targets When the Economy Keeps Growing. Wuppertal Institut für Klima, Umwelt, Energie GmbH, Berlin (2012)
32. Sorrell, S., Dimitropoulos, J., Sommerville, M.: Empirical estimates of the direct rebound effect: A review. Energy Policy. 37, 1356–1371 (2009)
33. Barker, T., Ekins, P., Foxon, T.: The macro-economic rebound effect and the UK economy. Energy Policy. 35, 4935–4946 (2007)
34. Hilty, L.M., Köhler, A., Von Schéele, F., Zah, R., Ruddy, T.: Rebound effects of progress in information technology. Poiesis & Praxis: International Journal of Technology Assessment and Ethics of Science. 4, 19–38 (2006)
35. Schriefl, E.: Steigender Energieverbrauch trotz verbesserter Energieeffizienz? Der „Rebound-Effekt“ und andere verbrauchstreibende Faktoren. Energiezukunft. 11, 128–133 (2008)
36. Gottron, F.: Energy Efficiency and the Rebound Effect: Does Increasing Efficiency Decrease Demand?. CRS Report for Congress, Washington, D.C. (2001)
37. Haas, R., Biermayr, P.: The rebound effect for space heating Empirical evidence from Austria. Energy Policy. 28, 403–410 (2000)
38. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly. 26, 13–23 (2002)
39. Berl, A., Gelenbe, E., Di Girolamo, M., Giuliani, G., De Meer, H., Dang, M.Q., Pentikousis, K.: Energy-Efficient Cloud Computing. The Computer Journal. 53, 1045–1051 (2010)
40. Mell, P., Grance, T.: The NIST Definition of Cloud Computing. NIST Special Publication 800-145, Gaithersburg, MD (2011)
41. Hilty, L.M.: Information Technology and Sustainability: Essays on the Relationship between Information Technology and Sustainable Development. Books on Demand, Norderstedt (2008)
42. Ritchey, T.: Fritz Zwicky, Morphologie and Policy Analysis. In: 16th EURO Conference on Operational Analysis. Brussels (1998)
43. Deitermann, M., Schmolke, S., Rückwart, W.-D.: Industriebuchführung mit Kosten- und Leistungsrechnung. Winklers, Braunschweig (2008)

Understanding Platform Loyalty in the Cloud: A Configurational View on ISV´s Costs and Benefits

Dominik Dellermann[1], Christopher Jud[2], and Fabian Reck[3]

[1] University of Kassel, Chair for Information Systems, Kassel, Germany
dellermann@uni-kassel.de
[2] University of Stuttgart, Chair for Information Systems, Stuttgart, Germany
jud@wius.bwi.uni-stuttgart.de
[3] University of Bamberg, Chair for Innovation Management, Bamberg, Germany
fabian.reck@uni-bamberg.de

Abstract. Platform-as-a-service (PaaS) providers are increasingly engaged in nurturing vibrant ecosystems of independent software vendors (ISVs) by offering standardized services. However, cloud ecosystems have also been known for its fluctuation and high rates of desertion. A currently under-researched explanation for this low traction and high rates of fluctuation may lie in the fact that ISVs face considerable costs when joining and acting on a specific platform. If these costs are too high, they can rapidly outweigh the additional value generated by the ecosystem. This study therefore explains the role of different configurations of cost-inducing factors and resource benefits in influencing an ISV´s platform loyalty. By using a configurational approach based on fuzzy-set qualitative comparative analysis (FsQCA), we display complex interactional effects of cost and benefits as causal conditions on ISVs' intention to stay in the ecosystem and thus provide valuable insights for both practice as well as theory on platform ecosystems.

Keywords: PaaS, cloud, ISV, transaction costs, FsQCA

1 Introduction

The ubiquitous provision of on-demand computing resources via the internet sets cloud computing in the spotlight of practitioners and researchers [1]. The fastest growing segment of this on-demand service delivery phenomenon is the platform-as-a-service market (PaaS). PaaS refers to an on-demand programming environment for developers of software applications, which promises to make software development more efficient [2]. IT industry leaders like Amazon Web Service (AWS) or Salesforce provide a scalable cloud platform, i.e. an expandable code base, which offers a large amount of independent third-party developers' important resources such as computing power and software databases to enable the development of applications that extend the basic functionality of the platform [3]. Such applications can be vice versa offered as software-as-a-service (SaaS) on the platform provider´s marketplace. By the numbers, Salesforce´s cloud infrastructure consists of approximately 5.5 million apps and more than 4 billion transactions a day.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Dellermann, D.; Jud, C.; Reck, F. (2017): Understanding Platform Loyalty in the Cloud: A Configurational View on ISV´s Costs and Benefits, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 514-528

This new logic of software development in the cloud has amplified the need for innovation from an ecosystem of third-party developers, called independent software vendors (ISVs) [4]. The focus of a PaaS provider is therefore to attract ISVs to join the ecosystem and facilitate innovation as well as the generation of complementary value propositions by offering standardized cloud services to an anonymous number of ISVs [5]. However, even famous examples like Microsoft´s cloud platform Azure show that it remains challenging to gain solid traction among ISVs. Furthermore, platform ecosystems have also been known for their fluctuation and high rates of desertion [6].

A currently under-researched explanation for disloyal behavior may lie in the fact that ISVs face considerable costs when joining and acting on a specific cloud platform. If these costs are too high, they can rapidly outweigh the additional value generated by the PaaS ecosystem [7] and induce ISVs' disloyal behavior and in its strongest manifestation the abandonment of a platform. Prior studies have mainly focused on the motivational factors and the relational rents that initially motivate ISVs to join an ecosystem [8, 9]. However, the costs associated with this choice and how particularly the interplay of the benefits and costs influence the decision to stay in a PaaS ecosystem has yet to receive research attention.

Previous studies that addressed this question were primarily focusing on a technological perspective and solely on the coordination costs related to platform dependencies [6] or the lock-in effects of cloud platforms. However, to provide a more holistic analysis of costs which affect sustainable traction among ISVs, also economic dimensions need to be included as entering service partnerships with a PaaS provider might induce a cost disadvantage relative to vertically integrated structures. However, managers are willing to accept certain costs if they expect superior returns and benefits [9,10].

We therefore propose that the interplay of transaction cost inducing hazards and resource-based benefits shape the decision of an ISV to stay loyal to a PaaS provider. As traditional regression-based net effect models are not capable to capture the complex interplay of different cost and benefit dimensions in shaping PaaS-related decision making, we apply configurational theory as theoretical lens and qualitative comparative analysis (QCA) as a way of analysis for understanding the complexities of PaaS ecosystems [11] and therefore address the following research question: *Which configurations of cost-inducing hazards and resource-based benefits maximize ISVs´ platform loyalty?*

To answer these questions, we analyze data from a survey of 42 ISVs on five leading cloud platforms by applying fuzzy set Qualitative Comparative Analysis (FsQCA) [12]. This case-oriented method enables us to analyze asymmetric and complex causal effects by extracting configurations that consistently lead to the platform loyalty [11, 13]. Our results reveal the role of cost-inducing hazards and resource benefits in influencing the loyalty of ISVs on certain cloud platforms. We therefore show that PaaS provider should balance such cost and benefit dimensions to gain solid traction among ISV and build sustainable ecosystems around their platform.

2 Conceptual Background

2.1. Platform Loyalty

A high number of complementary apps are central to the success of platform ecosystems [5]. Hence, PaaS providers are increasingly engaged in building vibrant ecosystems of ISVs around their platform [14]. For the long-term success and stability of a platform it is not only crucial to attract a large base of ISVs that produce high-quality applications but also to keep them loyal to the respective platform [15]. However, many platforms are plagued by high rates of platform desertion as ISVs stop developing applications [6], which represents the strongest manifestation of disloyalty. Further representations of disloyal behavior of ISVs might include factors like for instance showing reluctance to invest in the relation, stop curating existing applications on a specific platform or start multihoming [16]. Especially when ISVs change to rival platforms, the spillover of knowledge to competitors is a common threat to the PaaS provider that results from disloyal behavior [5]. Loyal behavior of a platform´s ISVs is a critical performance indicator. Furthermore, the ability of a PaaS provider to retain ISVs within their ecosystem is vital for success. Although previous work on developers´ loyalty on a platform in the B2C market also uses alternative explanations of platform loyalty [e.g. 17], we attempt to take an economic exchange lens, as such a perspective is especially important in the case of B2B applications. In this context, a PaaS provider should balance the effort required by the ISV to continue developing an application on the platform (i.e. cost-inducing factors) and the resource benefits offered to the ISV to stabilize the ecosystem and guarantee sustainability.

2.2 A Configurational Perspective on Platform-related Costs and Benefits

One core assumption of our study is that the influence of single factors on a specific outcome depends on the overall configuration of these factors rather than the net effects of each individual factor. Thereby, we abstain from conventional, rather reductionist variance models. Taking a configurational perspective on the conditions leading to platform loyalty is suitable for two reasons.

On the one hand, configurational approaches treat whole sets of elements as predictors rather than single variables [11]. These sets simultaneously explain the outcome of interest. Thus, one major advantage of configurational theory is its ability to explain synergetic and complementary effects [12]. This resonates well with current theoretical perspectives on the complexity of ecosystems as well as platform decisions [10]. For instance, ISVs are willing to accept cost if the benefits outweigh such. Hence, each variable in isolation may have a different effect on a decision than in combination with other elements.

On the other hand, configurations can display asymmetric relations between conditional and outcome variables rather than just symmetric ones [11]. So, configurational theory implies equifinality between different conditions and configurations [12]. Consequently, these conditions may either be sufficient or necessary causes of a dependent variable. According to both organizational [13] and information systems research [18] such notions possess superior accordance to organizational realities which are to a large part bounded to the larger context. As El-Sawy et al. [11] point out, this perspective thus particularly fits to explaining the ISV's decision to continue developing on a certain platform.

2.3. Research Framework

Fig. 1 illustrates the framework of our research. The framework comprises two facets of causal conditions for an ISV´s loyalty on a cloud platform. These two dimensions are theoretically grounded in TCT and the resource-based view. It therefore proposes that from the perspective of ISVs the configuration of four cost-inducing hazards (i.e. platform specificity; behavioral, market and technological uncertainty) and three resource benefits (i.e. technological, social and commercial capital) influences the intention of software vendors to stay in a certain PaaS arrangement.

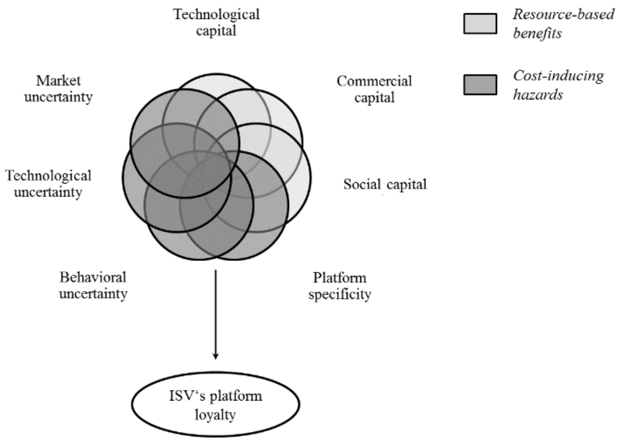


Figure 1. Research framework

The probably most prominent theoretical approach to explain boundary decisions associated with interfirm exchange (e.g. outsourcing decisions) is Transaction Cost Theory (TCT) [19 - 21]. TCT defines the costs of economic exchange (i.e. search and information costs; investments in social relations; opportunity costs) within the boundaries of a specific system (i.e. market or hierarchies) [22]. From TCT´s perspective developing applications on cloud platforms therefore might induce a cost disadvantage relative to vertically integrated structures of software development [7]. In the PaaS context, the required transactions for using an on-demand programming environment to develop software applications always create costs. According to TCT such costs mainly arise from two cost-inducing hazards: uncertainty and its sub-dimensions as well as specificity [23]. We therefore use these dimensions as cost predictors for platform loyalty.

Platform specificity: The specificity of a certain cloud platform represents the first hazard for ISVs. Platform specificity refers to the software migration between different PaaS providers [24] as well as the value of ISV´s assets within alternative PaaS relations [23]. For instance, cloud platforms require investments in relation-specific knowledge to participate in the ecosystem and capitalize from the access to complementary resources [20]. Apart from requirements to adapt software applications that are locked-in by specific APIs or for instance proprietary data storage implementations, high investments requirements create dependencies with the PaaS provider. This increases switching costs making it difficult for the ISVs to leave the actual cloud ecosystem and move to another platform [16].

Uncertainty represents the second cost-inducing hazard for ISVs. This TCT dimension is defined as the absence of complete information, which in turn leads to an ISV's inability to predict its surroundings accurately [25]. The uncertainty dimension can be subdivided in the volatility of market conditions (e.g. market, demand, and competitive environment), technological requirements (e.g. technological volatility) as well as the behavior (e.g. opportunism) of the PaaS provider [23].

Market uncertainty: Market conditions are crucial contextual conditions for ISVs. By developing complementary apps upon a platform, software vendors typically attempt to occupy a specific niche market [13]. Unpredictable user demand, substitute software products or changes in the competitive environment therefore increase the costs of ISVs.

Technological uncertainty: The second dimension of uncertainty in PaaS relations spans the difficulties to accurately predict the technological requirements. Particularly in the context of cloud platforms the quality of the service offering is hard to forecast. Furthermore, technological complexity and changes of specifications like for instance interfaces are frequently the most significant sources of uncertainty [7].

Behavioral uncertainty: Contrary to technological or market uncertainty, which is not directly related to the PaaS provider, behavioral uncertainty arises from the complexity of service performance evaluation. The platform provider for instance might follow its individual and act opportunistically [19]. ISVs therefore need to monitor the PaaS provider to detect opportunistic behavior like for instance exploiting resources or poaching in the ISV's niche [9].

From a resource-based perspective however, an ecosystem offers relational rents by providing the access to resources [26]. When ISVs decide to join an ecosystem, advantages or benefits need to be gained to make the commitment to a platform attractive [27, 28]. There are various forms of motives which drive ISVs to join and stay in an ecosystem. Most of them are related to getting access to a certain resource as technical or commercial resources. Kude et al. [9] grouped motives into three categories: technological capital, commercial capital and social capital.

Technological capital: By technological benefits we refer to access to technological resources that an ISV gains by joining a cloud ecosystem. These are for instance the PaaS provider's ability to supply integrated systems and its capability to innovate these systems [29]. First, the benefits from integrated systems arise for instance if the products and solutions offered by the ISVs offer only a small solution-space. Hence, this single solution gains value through the interoperability with a larger system of modules. Cloud ecosystems offer various technological resources like Applications Programming Interfaces (API), Software Development Kits (SDK) or Integrated Development Environments (IDE) [30]. Such resources allow the ISV to individually develop extending applications. Another benefit from the technological point of view is the availability of standards and technologies which are offered or hosted by the platform [9].

Commercial capital: Commercial capital refers to the PaaS provider´s or the whole cloud ecosystem´s marketing capabilities as well as its service and distribution networks. Especially, if ISVs have limited internal capabilities to heavily invest in marketing activities and to set up distribution networks, the access to such forms of commercial capital is crucial [31]. When partnering with huge platform providers like for instance Salesforce, ISVs can benefit from marketing and distribution capabilities to increase the awareness among many potential users by providing access to broad markets [32] or leverage highly visible cloud ecosystems to gain attention for own products. Distribution channels are for example AppStores. This kind of phenomenon also can be found in the domain of Enterprise Software where SAP or Microsoft offer respective features on their platforms. Using the possibilities of AppStores, ISVs can offer their solutions to end-users without the need to set up own distribution channels.

Social capital: Social capital refers to the PaaS provider´s reputation that is often aligned with the brand of this particular firm [32]. As ISVs are frequently unknown due to the limited reputation, customers might suspect the quality as well as reliability of their value propositions services. Nevertheless, such trustworthiness is a crucial topic in the software industry since the quality of SaaS offerings as well as the knowledge and experience of the vendor is difficult to assess a priori [9]. If a platform has reputation, e.g. AWS's reputation to produce and sell high-class products, ISVs may benefit from this reputation. A reputation can lead among others to a premium in prices. Furthermore, social benefits include the use of communities where ISVs can exchange information with other ISVs or ask for support in case of problems during the development of applications [17].

3 Research Methodology

3.1 Fuzzy-set QCA

Data analysis was done via FsQCA. This set-theoretic approach emphasizes the effects of the whole rather than its pieces. Hence, the method explicitly acknowledges that research cases are multidimensional. FsQCA thereby draws on measures of consistency and coverage to evaluate the predictive power of single configurations towards the outcome of interest. The first indicator, consistency, is analogous to correlation estimates in statistical methods. This value's meaning is the degree to which cases of a certain configuration agree in leading to a given outcome [12]. The second indicator, coverage, displays the degree to which a configuration accounts for the instances of an outcome. Hence, these values are analogous to R-square in regression analysis. FsQCA detects configurations with adequate consistency and coverage values in three steps which are: a) calibration, b) construction of truth tables, and c) truth table analysis [12].

Calibration represents the first step. This step is necessary because set-theoretic analysis is based on the degree of memberships of cases in a certain set of conditions (here, e.g. membership in the group of firms with high social capital). Thus, to obtain such so called fuzzy set membership scores all construct measures must be transformed to a scale ranging between 0 and 1 with 0 indicating full non-membership, 1 indicating full membership and 0.5 representing the crossover point [33]. Analogous to the calibration approach by Fiss [13], the observed maximum and minimum values within our sample for all variables specify full membership and full non-membership. The calculated scale midpoint (median of observed values) is the cross-over point. The fuzzy set memberships scores for each case were calculated via the calibration procedure in the FsQCA software program (version 2.5 [34]), with the three above mentioned values as calibration benchmarks.

The *construction and refinement of truth tables* represents the second step of analysis. In this context, a truth table is a matrix of all possible configurations of predictor conditions (in our case, 32 rows; in general, 2k, where k is the number of conditions observed [12]). Subsequently, the requirements of FsQCA demand a refinement of this truth table. The refinement procedure draws on two criteria to asses each possible configuration: frequency and consistency. The frequency criterion captures if and how many empirical cases exist which have a membership score of more than 0.5 and thus display membership in the configuration of interest more than non-membership. The standard threshold for frequency in medium-sized samples is 1 meaning that every configuration that exists in the empirical dataset will be part of the analysis [33]. The consistency criterion captures if a truth table row consistently leads to an outcome. This value should outreach .8, at least [12]. We choose a rather conservative threshold of .9 within this work.

In the third and final step, FsQCA analyzes the refined truth tables based on Boolean algebra, respectively counterfactual analysis, to be more specific. This step of analysis draws on the Quine-McCluskey algorithm which strips away factors that are inconsistently present or absent concerning the outcome [13]. By doing so, the algorithm excludes conditions that are no essential part of a sufficient configuration for the respective outcome. The result of this procedure encompasses two distinct solutions: the parsimonious solution and the intermediate solution. In the parsimonious solution, all simplifying assumptions derived from counterfactuals are included. In contrast, the intermediate solution only includes simplifying assumptions based on easy counterfactuals. Because of this dual algorithm, the intermediate solution necessarily represents a subset of the parsimonious solution, so that all conditions appearing in the parsimonious solution appear in the intermediate solution but not vice versa [12]. If a condition appears in the parsimonious solution, it passed a more thorough reduction procedure. In other words, the data provides particularly strong empirical evidence for the causality in this case. This condition thus displays a causal core of a configuration, while the periphery includes all conditions present in the intermediate solution [13].

3.2 Data Collection and Sample Description

Our sample consists of 750 firms equally distributed and randomly drawn among the ISVs of five market leading cloud platforms (i.e. Microsoft Azure, Oracle Cloud Platform, Amazon Web Services, SAP HANA, and Salesforce Force.com). There were two reasons for choosing exactly these platforms. First, all are instituted by established players and have shown an adequate amount of traction of their ISVs. Therefore, they can provide sufficient benefits for ISVs. Second, due to their high level of power imbalance they are perfectly suited for analyzing asymmetric third-party relationship and the corresponding costs related to that imbalance.

A web-crawling approach randomly collected sales contact data from the platforms´ app stores. This approach is congruent with previous surveys of third-party innovators [30]. Recipients were asked to forward the questionnaire to high-level executives (C-level; IT executives) as key informants who completed an online questionnaire containing the constructs of interest. We furthermore ensured confidentiality and anonymity to the participants.

In total, we obtained N=42 valid cases in which the data was complete. The resulting response rate of 5.6% is common in such settings. Non-response bias might still be an issue, so we compared responses of early and late respondents [35]. T-tests not reveal any significant differences ($p > 0.05$), hence we are confident to reject the presence of non-response bias in our study.

ISVs in our study were distributed among all five platforms (Microsoft Azure: 9; Oracle Cloud Platform: 4; Amazon Web Services: 2; SAP HANA: 9; and Salesforce Force.com: 18). Most of our respondents were high-level executives (C-level: 71.4%; BU executives: 19%). Participants in our sample indicated that they are highly experience in this topic (>10 years: 83.3%) and were experts in the context of our survey (95.2%).

3.3 Measurement Validation

To ensure validity we applied measures that were already developed and validated in prior studies in TCT [36 - 38]. For the benefit dimension, we developed scales based on the constructs of [9]. When necessary, we adapted scales slightly to the platform context. All items were rated on seven-point Likert scales. Through a pilot study with managers in the software industry, we pretested and refined our measurement instrument to ensure that each of them was clearly and unambiguously phrased.

Construct	Number of Items	MEAN	Standard Deviation	Loadings Range	Cronbach's Alpha	Composite Reliability	AVE
Outcome Variable							
Platform loyalty Adapted from [17]	3	5.65	1.36	.974	.941	.974	.949
Causal Conditions							
Platform specificity Adapted from [36]	4	4.99	1.41	.725-.931	.861	.907	.711
Market uncertainty Adapted from [37, 38]	4	4.61	1.62	.856-.952	.916	.941	.799
Technological uncertainty Adapted from [37, 38]	4	3.83	1.69	.785-.963	.925	.942	.804
Behavioral uncertainty Adapted from [37]	4	3.77	1.65	.848-.963	.894	.919	.743
Technological capital Following [9]	3	5.99	.97	.728-.954	.855	.900	.774
Commercial capital Following [9]	3	5.61	1.40	.796-.911	.834	.910	.751
Social capital Following [9]	3	5,62	1.36	.784-.941	.852	.912	.776

Table 1. Construct measures

We assessed reliability, convergent validity (see Table 1) to ensure that all values exceed the recommended threshold, supporting the notion of scale reliability. Furthermore, the AVE's square root exceed the shared variance between a single construct and all other constructs within model, reflecting discriminant validity [39]

We additionally conducted Harman's one-factor test [40] to assess for common method bias. The unrotated factor solution resulted in 5 factors explaining 80% of the variance (31 percent was the largest variance explained by one factor). Thus, common method bias is unlikely to be a problem.

4 Results

The results of the FsQCA reveal several patterns that explain how different configurations of cost-inducing hazards and resource-based benefits result in high or low levels of ISV´s loyalty. High levels of loyalty indicate a high intention to stay in the ecosystem while low loyalty indicates a high intention to leave. We extracted this pattern by comparing structures of different configurations [13]. Figure 2 shows the configurations resulting from FsQCA. Black circles indicate the presence of a condition, crossed-out circles indicate the absence of a condition, large circles indicate core condition, and small circles indicate peripheral conditions. Blank spaces indicate a condition may be either present or absent.

	Solutions for intention to stay					Solutions for intention to leave				
	1a	1b	1c	1d	1e	1	2	3a	3b	3c
Hazards										
Platform specifity		●	●	⊗	●	⊗	●	⊗	⊗	●
Behavioral uncertainty	⊗		●	⊗	●		⬤	⬤	⬤	⬤
Technological uncertainty	⊗	●		⊗	⊗	●	⊗		⊗	●
Market uncertainty	⊗	●	●	●	●	⊗	●	⊗	●	●
Benefits										
Technological capital	⬤	⬤	⬤	⬤	⬤	⨂	●	⨂	⨂	⨂
Commercial capital	●	●	●	⊗	⊗	⊗	⨂	⨂	⨂	⨂
Social capital	●	●	●	⊗	⊗	⊗	⊗	⊗	⊗	⊗
Consistency	.99	.92	.91	.95	.91	.95	.97	.94	.99	.94
Raw coverage	.47	.44	.39	.28	.23	.43	.32	.46	.38	.39
Unique coverage	.18	.04	.01	.03	.01	.02	.04	.06	.01	.02
Overall solution consistency			.91					.94		
Overall solution coverage			.73					.69		

Notes: Black circles indicate the presence of a condition, circles with "x" indicate its absence. Large circles indicate core conditions, small ones indicate peripheral conditions. Blank spaces indicate may be either present or absent.

Figure 2. Solutions for high and low platform loyalty

4.1 Configurations for High Platform Loyalty

We identified five different configurations that result in a high intention of the ISV to stay, which underlines the equifinality of achieving a certain outcome. Consistency for configurations ranges from 0.91 to 0.99 representing the degree to which these configurations of causal conditions result in high platform loyalty. Raw coverage, which describes the importance of a certain configuration in explaining the intended outcome, range from 0.23 to 0.47. The overall solution consistency shows these five solutions can consistently result in high platform loyalty with 89%. The overall solution coverage indicates that the extent to which these seven configurations cover high likelihood of risk cases is 76%. We compared the five configurations that result in high platform loyalty to extract four strong patterns:

- **I.** In platform ecosystems that offer high benefits and low levels of behavioral and environmental uncertainty, ISVs have a high level of platform loyalty no matter if the amount of investment in platform-specific resources is high or low (1a).
- **II.** Technological capital is a necessary condition for ISVs to stay in an ecosystem. Its presence is required to create platform loyalty (1a,b,c,d,e).
- **III.** ISVs are willing to accept all forms of uncertainty as well to heavily invest in platform-specific resources when all three benefits are prevalent (1b,c).
- **IV.** If the ecosystem does not offer social and commercial capital to the ISV, the combination of technological capital and technological stability is required (1d,e).

4.2 Configurations for Low Platform Loyalty

Furthermore, we identified five distinctive configurations that exceed minimum consistency threshold and result in a low level of platform loyalty. Consistency for configurations ranges from 0.94 to 0.99. Raw coverage ranges from 0.32 to 0.46. These five solutions can consistently result in low platform loyalty with 94% and cover 69% of cases with this outcome. Comparing the five configurations reveals three further important patterns:

- **V.** In cases where technological uncertainty is high and ISVs are not provided with sufficient technological capital by the platform, their loyalty diminishes (1)
- **VI.** Especially when behavioral uncertainty is high and commercial capital is not provided by the ecosystem the ISVs´ loyalty to the ecosystem suffers (2; 3a,b,c).
- **VII.** Although technological capital is present, high behavioral uncertainty, market uncertainty and platform-specificity outweigh technological benefits resulting in a low level of platform loyalty (2).

4.3. The Role of Cost-inducing Hazards and Resource-based Benefits in Maximizing Platform Loyalty

Out of these seven patterns derived from the comparison of configurations that lead to high and low platform loyalty, we can reveal holistic insights of the interplay of cost-inducing factors and resource benefits in influencing the intention of a third-party developer to stay in the ecosystem or to leave it. Based on the commonalities among the pattern we identified three holistic findings to explain the loyalty of complementors.

First, technological capital is a necessary condition for ensuring platform loyalty among the ISV's. In all our sample cases, which displayed high levels of loyalty, the ISV perceived access to valuable technological resources which help it provide and innovate its systems and software products. The results strongly indicate that without the superiority of the technological capital provided via the ecosystem, gaining strong traction among ISVs is hardly likely. Thus, platform owners must ensure ISV's access to technological resources to ensure their loyalty.

Second, technological still does not suffice for predicting platform loyalty. In situations of high behavioral uncertainty, volatile market environments, highly specific needs of platform-related investments and the absence of further benefits like commercial and social capital, ISV may display low levels of loyalty even if they can access technological capital. Hence, the presence of technological capital is a necessary but not sufficient condition for high levels of platform loyalty.

Still, combinations of technological capital with other cost- or benefit-related factors prove to be collectively sufficient [33] for ISV's loyalty. Excluding the rather trivial solution 1a covering cases in which ISV perceiving high benefits with low hazards from this discussion, the results demonstrate two main paths to high platform loyalty (Patterns III and IV). One the one hand, the absence of technological uncertainty in combination with technological capital is sufficient for ISV's loyalty (solutions 1d,e). This finding highlights the importance of the appropriation of resources over access to them [41]. If technological specifications in the platform remain stable, profiting from accessed technological capital is much easier so that ISVs can obtain value from the ecosystem. On the other hand, in situations of high uncertainty and highly specific investments to the platform, technological capital does not suffice alone (solutions 1b,c). In these cases, ISV's loyalty is only secured, if they are also provided with commercial and social capital. Hence, in situations where profit from technological capital is not guaranteed, ISVs want further benefits from the ecosystem. As there are circumstances under which technological capital both alone and in combination with the other two benefits leads to high loyalty, these two paths are equifinal.

Third, in contrast to technological capital, commercial and social capital seem to be of different importance. The solution table for low platform loyalty demonstrates that for low levels of loyalty, commercial and social capital are absent. As a converse argument, the presence of both commercial and social capital is a sufficient condition for high platform loyalty. Hence, commercial and social capital can enhance technological capitals effects for platform loyalty but seem to be rather substitutable add-ons in the eyes of ISVs.

Fourth and finally, while cost-inducing hazards play a role in predicting ISV's loyalty towards an ecosystem, there is none which cannot be overcome by the provision of technological, commercial and social capital. Behavioral uncertainty seems to play a quite important role. In four of the five paths leading to low levels of loyalty, the ISV sees the platform owner as a key source of risk. This result highlights the outstanding importance of the platform owner in ensuring ISVs traction to the ecosystem. The other three hazards represent rather peripheral conditions which are partially interchangeable concerning their effect on platform loyalty. The permutations of the main solutions for high (1a,b,c,d,e) and low platform loyalty (3a,b,c) stem from different merely inconsistent combinations of these. So, these factors play rather minor roles in causally explaining platform loyalty.

5 Discussion and Conclusion

Our study describes the interplay of cost-inducing hazards and resource benefits of PaaS offerings in explaining platform loyalty. By comparing different configurations that result in high and low platform loyalty, we identified seven patterns that describe the role of cost-inducing hazards and resource benefits in shaping ISVs´ platform loyalty. By drawing on configurational theory and applying FsQCA we can provide a much more fine-grained perspective on the complex causality associated with our dependent variable, platform loyalty. In doing so, we contribute to theory on platform ecosystems and ISVs loyalty in various ways.

First, we introduce the concept of hazards related to transaction costs to this phenomenon and evaluate their role in weakening or strengthening ISVs loyalty. Thereby, we find that especially behavioral uncertainty serves as a key cause of low levels of loyalty while the other hazards play a rather peripheral role. Still, also behavioral uncertainty is accompanied by the absence of at least two of the resource benefits, so that only in combination with them behavioral uncertainty is a sufficient cause of disloyalty. These two insights advance theories on platform loyalty by a) outlining the platform owner's behavior as a key driver of platform-related costs in addition to previously identified cost drivers such as e.g. the management of technological dependencies [6] and b) indicating that while also highly loyal ISVs may face high levels of hazard and consequentially high transaction costs access to valuable resources outweighs these costs. Hence, our findings support a notion of ISVs loyalty to be driven by opportunity-seeking rather than risk-avoiding motives. This gives hints that the issue ISVs loyalty to platform ecosystems might hold important contextual differences compared to traditional business partnerships like for instance between outsourcing contractors [20, 21] or service providers and customers [42]. Further research might specify these differences and examine why these exist.

Second, drawing on configurational theory and applying FsQCA allows us to capture asymmetrical causal relationships and thus shed light on how the effects of distinct resource-based benefits on ISVs platform loyalty differ. Thereby, we provide a more differentiated view than previous studies [e.g. 3, 8, 9, 42] on how resource benefits motivate platform loyalty. Our analysis reveals that technological capital is a necessary condition for high loyalty, while commercial and social capital are collectively sufficient conditions. In other words, platform owners do not remain loyal if they do not perceive clear benefits on the technology side but were loyal in all the cases within our sample where all technological, commercial and social capital were high. These findings indicate a preference hierarchy of ISV. They may join and stay on the platform mainly with the goal of enhancing their own technological capabilities by the architectural features the software platform offers. If the platform's technological features are stable, the ISV will stay loyal even if cost-inducing hazards are high. Commercial and social capital are rather valued if such stability is not given and may partially compensate in this case. This fact challenges the findings of previous studies that were focusing on network effects as the main (and sometimes only) driver of platform traction [43, 44] We suppose future research on ecosystems to analyze why ISVs are rather technology-driven. Furthermore, it would be interesting to investigate if these perceptions and preferences match with value opportunities in ecosystems or rather indicate strategical "blind spots" of the ISVs which hinder them from benefiting more from commercial resources and network linkages gained through platform membership.

On the practical side, implications of our research show the impact of different influencing factors on if ISVs remain on a platform or leave. This is particularly important for offering cloud services like PaaS, which are rather addressing a mass market then providing customized service value proposition. We therefore provide insights on how such standardized PaaS offerings should balance cost-inducing hazards and benefits to gain solid traction among a huge number of anonymous ISVs. The different configurations derived from FsQCA may serve as a blueprint for PaaS providers in designing their ecosystems of third-party software development.

However, this work is not without limitations. Although FsQCA is particularly suitable for medium sample size (n=5-50) an increased sample size probably cloud enhance the insights and generalizability of our results. Second, we did not include platform characteristics in the analysis of our empirical data so far. However, the characteristics and the explicit value proposition of the single PaaS providers vary. For instance, SAP HANA´s in memory data base is a very specific service offering compared to the other platforms in our sample. We therefore intend to include such platform characteristics in the further progress of our study. Third, this study includes a potential selection bias as the respondents of the survey that are still on the platform might over represent loyal ISVs. To proceed this study, attempt should therefore point towards gathering data from ISVs, who already left the platform, to gain more robust results.

References

1. Armbrust, M., A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A View of Cloud Computing", Communications of the ACM, 53(4), pp. 50–58.
2. Leimeister, S., M. Böhm, C. Riedl, and H. Krcmar, "The Business Perspective of Cloud Computing: Actors, Roles and Value Networks", ECIS Proceedings, 2010, Paper 56.

3. Tiwana, A., B. Konsynski, and A.A. Bush, "Research Commentary: Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics", Information Systems Research, 21(4), 2010, pp. 675–687.
4. Ghazawneh, A. and O. Henfridsson, "Balancing Platform Control and External Contribution in Third-party Development: The Boundary Resources Model", Information Systems Journal, 23(2), 2013, pp. 173–192.
5. Boudreau, K.J., "Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software App Developers and Patterns of Innovation", Organization Science, 23(5), 2012, pp. 1409–1427.
6. Tiwana, A., "Platform Desertion by App Developers", Journal of Management Information Systems, 32(4), 2015, pp. 40–77.
7. Williamson, P.J. and A. de Meyer, "Ecosystem Advantage", California Management Review, 55(1), 2012, pp. 24–46.
8. Ceccagnoli, M., C. Forman, P. Huang, and D.J. Wu, "Co-creation of Value in a Platform Ecosystem: The Case of Enterprise Software", MIS Quarterly, 36(1), 2012, pp. 263–290.
9. Kude, T., J. Dibbern, and A. Heinzl, "Why Do Complementors Participate? An Analysis of Partnership Networks in the Enterprise Software Industry", IEEE Transactions on Engineering Management, 59(2), 2012, pp. 250–265.
10. Chiles, T.H. and J.F. McMackin, "Integrating Variable Risk Preferences, Trust, and Transaction Cost Economics", Academy of Management Review, 21(1), 1996, pp. 73–99.
11. El Sawy, O.A., A. Malhotra, Y. Park, and P.A. Pavlou, "Research Commentary: Seeking the Configurations of Digital Ecodynamics: It Takes Three to Tango", Information Systems Research, 21(4), 2010, pp. 835–848.
12. Ragin, C.C., "Redesigning Social Inquiry: Fuzzy Sets and Beyond", Wiley Online Library, 2008.
13. Fiss, P.C., "Building better causal theories: A fuzzy set approach to typologies in organization research", Academy of Management Journal, 54(2), 2011, pp. 393–420.
14. Wareham, J., P.B. Fox, and J.L. Cano Giner, "Technology Ecosystem Governance", Organization Science, 25(4), 2014, pp. 1195–1215.
15. Goldbach, T. and A. Benlian, "How Social Capital Facilitates Clan Control on Software Platforms to Enhance App-Developers' Performance and Success", ICIS Proceedings, 2015.
16. Kude, T. and J. Dibbern, "Tight versus Loose Organizational Coupling within Inter-firm Networks in the Enterprise Software Industry – The Perspective of Complementors", AMCIS 2009 Proceedings, 2009, p. 666.
17. Goldbach, T., V. Kemper, and A. Benlian, "Mobile Application Quality and Platform Stickiness under Formal vs. Self-Control—Evidence from an Experimental Study", ICIS Proceedings, 2014.
18. Liu, Y., J. Mezei, V. Kostakos, and H. Li, "Applying Configurational Analysis to IS Behavioral Research: A Methodological Alternative for Modelling Combinatorial Complexities", Information Systems Journal, 2015.
19. Williamson, O.E., "The Economic Intstitutions of Capitalism", Simon and Schuster, 1985.
20. Aubert, B.A., S. Rivard, and M. Patry, "A Transaction Cost Model of IT Outsourcing", Information & Management, 41(7), 2004, pp. 921–932.
21. Watjatrakul, B., "Determinants of IS Sourcing Decisions: A Comparative Study of Transaction Cost Theory Versus the Resource-based View", The Journal of Strategic Information Systems, 14(4), 2005, pp. 389–415.
22. Ngwenyama, O.K. and N. Bryson, "Making the Information Systems Outsourcing Decision: A Transaction Cost Approach to Analyzing Outsourcing Decision Problems", European Journal of Operational Research, 115(2), 1999, pp. 351–367.

23. Rindfleisch, A. and J.B. Heide, "Transaction Cost Analysis: Past, Present, and Future Applications", Journal of Marketing, 1997, pp. 30–54.
24. Baldwin, C.Y. and K.B. Clark, "Design rules: The Power of Modularity", MIT press, 2000.
25. Milliken, F.J., "Three Types of Perceived Uncertainty about the Environment: State, Effect, and Response Uncertainty", Academy of Management Review, 12(1), 1987, pp. 133–143.
26. Dyer, J.H. and H. Singh, "The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage", Academy of Management Review, 23(4), 1998, pp. 660–679.
27. Evans, D.S., A. Hagiu, and R. Schmalensee, "Invisible Engines: How Software Platforms Drive Innovation and Transform Industries", MIT press, 2008.
28. Smedlund, A., "Digital Health Platform Complementor Motives and Effectual Reasoning", IEEE, 2016.
29. Hagedoorn, J., "Understanding the Rationale of Strategic Technology Partnering: Interorganizational Modes of Cooperation and Sectoral Differences", Strategic Management Journal, 14(5), 1993, pp. 371–385.
30. Benlian, A., D. Hilkert, and T. Hess, "How Open is this Platform? The Meaning and Measurement of Platform Openness from the Complementors' Perspective", Journal of Information Technology, 30(3), 2015, pp. 209–228.
31. Rao, P.M. and J.A. Klein, "Growing Importance of Marketing Strategies for the Software Industry", Industrial Marketing Management, 23(1), 1994, pp. 29–37.
32. Ahuja, G., "The Duality of Collaboration: Inducements and Opportunities in the Formation of Interfirm Linkages", Strategic Management Journal, 21(3), 2000, pp. 317–343.
33. Schneider, C.Q. and C. Wagemann, "Set-theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis", Cambridge University Press, 2012.
34. Ragin, C.C., K.A. Drass, and S. Davey, "Fuzzy-set Qualitative Comparative Analysis 2.0", Tucson, Arizona: Department of Sociology, University of Arizona, 2006.
35. Armstrong, J.S. and T.S. Overton, "Estimating Nonresponse Bias in Mail Surveys", Journal of Marketing Research, 1977, pp. 396–402.
36. Heide, J.B. and G. John, "Alliances in Industrial Purchasing: The Determinants of Joint Action in Buyer-Supplier Relationships", Journal of Marketing Research, 1990, pp. 24–36.
37. Stump, R.L. and J.B. Heide, "Controlling Supplier Opportunism in Industrial Relationships", Journal of Marketing Research, 1996, pp. 431–441.
38. Walker, G. and D. Weber, "A Transaction Cost Approach to Make-or-buy Decisions", Administrative Science Quarterly, 1984, pp. 373–391
39. Fornell, C. and D.F. Larcker, "Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics", Journal of Marketing Research, 1981, pp. 382–388.
40. Podsakoff, P.M., S.B. MacKenzie, J.-Y. Lee, and N.P. Podsakoff, "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies", Journal of Applied Psychology, 88(5), 2003, p. 879.
41. Tee, R., and A. Gawer, "Industry architecture as a determinant of successful platform strategies: A case study of the i-mode mobile Internet service", European Management Review, 6(4), 2009, pp.217–232.
42. Lee, M. and L.F. Cunningham, "A Cost/Benefit Approach to Understand Service Loyalty", Journal of Service Marketing, 15(2), pp. 113–130.
43. Farrell, J. and P. Klemperer, "Coordination and lock-in: Competition with switching costs and network effects", Handbook of industrial organization, 3, 2007, pp. 1967–2072.
44. Katz, M.L. and C. Shapiro, "Systems competition and network effects", The journal of economic perspectives, 8(2), 1994, pp. 93–115.

Enterprise System Renewal – The Divergence Between Perception and Reality

Phillip Haake[1], Silvia Schacht[2], Benjamin Mueller[3], and Alexander Maedche[2]

[1] University of Mannheim, Institute for Enterprise Systems (InES), Mannheim, Germany
{haake}@es.uni-mannheim.de
[2] Karlsruhe Institute of Technology (KIT), Institute of Information Systems and Marketing (IISM) and Karlsruhe Service Research Institute (KSRI), Karlsruhe, Germany
{silvia.schacht,alexander.maedche}@kit.edu
[3] University of Groningen, Faculty of Economics and Business, Groningen, Netherlands
{b.mueller}@rug.nl

Abstract. E-commerce based companies rely on the effective use of the information systems used to support their processes. Accordingly, managers place a great emphasis on the success of projects to introduce such systems. However, research increasingly suggests that project success may not be as objective as one would assume or hope. Quite contrary, as our work will show, project success is often constructed by the stakeholders involved in the project. Extending prior research, we investigate how different groups of stakeholders construct their own perception of project success and how these different perceptions influence each other. Through our work, we provide management with insights into threats to a reliable project management approach for critical IS projects and identify a few major drivers that need to be accounted for to make sure that such critical projects really are successful.

Keywords: Renewal Project, Enterprise System, Project Success, Case Study, Subjectivity of Success

1 Introduction

Issues with backend enterprise systems (ES) can have a critical impact on business performance. In particular, e-commerce based companies suffer from an ineffective usage of backend ES and might struggle to cope with competition. A prominent example for the effect of such issues on once prosperous businesses is the Otto Group, a large German distance retailer [1]. Due to under-investments, the technology base of the Otto Group had become outdated and scattered. For instance, 130 different ES were used to support the frontend services the customers interacted with. This led to complicated and delayed technological changes and made internal processes inefficient. In turn, the inefficient processes and the high complexity of the backend ES affected the number of available products online and the lapse rate at Otto Group. In sum, Otto has failed to reign in that toxic complexity and to manage a successful renewal project.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Haake, P.; Schacht, S.; Mueller, B.; Maedche, A. (2017): Enterprise System Renewal - The Divergence Between Perception and Reality, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 529-543

Cases like this sparked our interest in the reasons for successful ES renewal projects. We were able to select a multi-channel fashion retailer in Central Europe as our case site. In particular, we were enabled to analyze a renewal project [2] for a Product Information Management System (PIMS) in the e-commerce department of a multi-channel retailer. IS project success research in general, which is relevant for this type of project analysis, can be subdivided into two main streams. In the first stream, researchers are assessing project management success by the 'Iron Triangle' of a project's cost, time, and quality [3, 4]. Thereby, researchers aim to identify indicators that enable projects to reduce costs and time and increase the quality of the resulting product. The second stream focusses on the success of the project's product [5]. More recently, these aspects have been combined and resultant customer satisfaction [6] has become the focus. Prior research with a specific focus on ES implementation and renewal projects has been based on the assumption that such projects are essentially disruptive and lead to changes to the technology as well as the work environment and the task [7]. However, we have found evidence in our exploratory case study that even scheduled events that have no influence on the task can cause significant disruptions and adaptation efforts. As project success can also be considered as socially constructed and perceived [8], the investigation of the discrepancy between perceived success and the reality is necessary to come to a real understanding of project success. Thus, we aim to explain the diversion between perceived and actual success of an ES renewal project. We applied the critical realism perspective to identify the mechanisms behind the development of such a diversion. For our research, this is the appropriate approach, because it allows to focus on establishing causality [9]. On this basis, we formulate the following research question: *How and why is there a discrepancy between end-users' perceived and real ES renewal project success?*

We aim to provide an overview of the mechanisms that are behind the different perceptions of the renewal project. This will result in a type II theory [10] and several relevant practical implications. In the following section, we provide a brief overview on the theoretical background of our research. We discuss the methodology of our explanatory, longitudinal single case study in section three, before presenting and discussing our findings in section four and section five, respectively. Finally, we conclude our research by summarizing the key results, discussing the limitations and contributions of our study, and providing an outlook on future research.

2 Theoretical Background

There is a lot of research about IS projects and system success as well as failure [8, 11–13]. However, there is no agreed definition of IS project success in general [5]. In part, this is attributable to the multiple facets of project success. As success is a multi-dimensional construct, it is subjective and depends on perceptions [5]. A distinction between overall project success and project management success is an important step in the analysis process for determining whether a project should be considered a success or a failure [14]. Project management success is often measured as adherence to planning in the form of the 'Iron Triangle' of time, budget, and quality. The Iron Triangle is popular in research [3, 15] and in practice [5]. On the other hand, there is also the product related dimension of the project outcome, which can be coined product success [5]. As an example, we look at the definition by Basten, Joosten and Mellis [6]. In their definition, the adherence to planning in the form of the Iron Triangle is the definition of project management success. Product success is defined as the effect of the product in terms of organizational benefits or customer satisfaction. Overall project success is determined as the successful combination of these two success dimensions.

Research on IS project success is generally based on the measurement of IS product success [16, 17]. The wide range of criteria for the measurement of product success, make it an ambiguous measure. Aspects directly related to IS products can be used for an assessment based on system performance, for instance, perceived usefulness, information quality, and system quality (e.g. ease of use) [18] or system reliability [15]. A more general approach to IS project success in the past focused on the combination of project management and a successful product [3]. However, the perception of information system failure or success is largely stakeholder-dependent [19]. IS project success is, therefore, socially constructed and perceived [8]. This understanding has given rise to a performative perspective on IS project success [8]. A performative perspective is based on the identification of actors and their relational effects in the networks of IS projects. Actors in a project network value different aspects of an implemented IS or of a project differently. Consequently, the actors measure and evaluate IS project success differently. This aspect is crucially related to organizational sensemaking. Organizational sensemaking is focused on determining what an event means for members of an organization [20]. Sensemaking is based on the idea of retrospectively making sense of events [20] such as a renewal project. During the course of such a project, sensemaking in a group can be influenced by the social dynamics in the group of affected people. For instance, it is a crucial characteristic of a good team that members show a great deal of synergy and loyalty to each other and to their leader [21, 22]. However, these are also factors, which can lead to groupthink [21]. In particular, hierarchical groupthink, which originates in the desire of individuals to please their leader by agreement in opinion, can have a strong influence on the assessment of project success. Especially, since employees' sensemaking can be strongly influenced by a management's narrative [22]. For instance, employees (i.e. end-users) develop a reliable system [23] to cope with perceived adversity, which might be caused by technological glitches in their work environment. Whether overcoming the situation as a group can give them a collective mind and feeling of success is the subject of further research.

3 Methodology

To answer the aforementioned research question, we decided to analyze one case company longitudinally with a single case study approach. We reviewed the transition and change of end-users' expectations in the organizational context of the e-commerce subunit, which is the unit of analysis. Thereby, we aim to explain the deviation of perceived project success over time. This single instant serves as a starting point for the search for an explanation [24]. In combination with insights from the literature, this holistic view allows us to develop the explanation [10] presented at the end of this paper.

3.1 Case Description

We acquired a project for the analysis of the renewal and adaptation of an ES. During the course of the single case study, we analyzed the development of perceived and real ES renewal project success. As case company, we selected a multi-channel fashion retailer with a sizable online shop, which is located in Central Europe. Thereby, the e-commerce department (in the following referred to as FASHION) and its Product Information Management System (PIMS) were at the center of our research. A PIMS allows to centrally manage all information required to market and sell products on distribution channels such as FASHION's online shop and marketplaces. FASHION is a department of two managers, content managers, and supporting technicians. FASHION's deputy department head characterizes his business unit in the following way: *"I see us as a hub which compresses the product information and provides access to sales channels [for other departments in the company]." Due* to changing requirements, FASHION regularly undergoes changes of its e-commerce platform. At the center of the change process, the new PIMS release was supposed to significantly improve PIMS overall and the Web-Client version in particular. The release was supposed to update the software to the originally contracted level, since this version had not been ready for renewal for the original project. Changes in roles or assigned tasks were not planned. At the time, FASHION had a lead and a deputy technician who were responsible for the online-shop system and PIMS, which were the relevant IS for e-commerce. The deputy technician had started his new job a month before the introduction of the new release.

There were 84 recorded users of the PIMS Web-client, which include the department heads, their deputies, the content management team, and users in various purchasing departments. We only evaluate the PIMS Web-Client, which is a content management system for product information, classification in the structure of the online-shop, and management of product images. Content managers focus on texting and classification of products. Texting and classifying a product took on average 7-8 minutes before the renewal project. Up to 50 articles had to be processed by a content manager per day. FASHION employs two teams of four content managers and two interns. The other employees in the purchasing department mainly search and read in the PIMS. Team leads in the content management team use a Master client version of the PIMS, which allows them to assign work packages of texting and classification work to team members. Two months after the renewal of the new release of the PIMS, one of the two remained team lead for the texting group and the other became head of a newly created product image production team.

3.2 Data Collection

We used several data collection methods during the case study. Our data collection included 22 *semi-structured interviews*, *participant observation*, and *document analyses*. Semi-structured interviews are defined as interviews in which pre-formulated questions are used, but not strictly adhered. New questions can emerge during the conversation [25]. We interviewed different user types, such as content managers, team leaders, managers, and employees of the technical support. The multi-level analysis in our research made it necessary to include different user categories for the analysis of the specific ES [26]. Thereby, we aimed to get an integrated view of user adaptation and developing experiences of the renewal initiative by interviewing a carefully selected set of people over the course of the renewal project for eight months. We interviewed as many individual users of the PIMS as necessary to get an understanding of the typical user role in FASHION. Interviews with management focused on the department head and his deputy who were responsible for the PIMS project. Technology support included the positions responsible for the e-commerce related IT services and those responsible for the particular IS project. The first author conducted the interviews in person, recorded and transcribed them. The interviews lasted typically between 45 to 60 minutes. Before an interview, we provided some information to the interviewees regarding the interviewer, the background and purpose of the study, and the anonymity and use of gathered data [27]. We conducted the semi-structured interviews at three points of time: (1) before the start of the project, (2) shortly after the renewal, and (3) after employees had settled in with the new system (see Figure 1). This time frame was chosen because researchers suggest a gap of one month between perception of a new system and usage measurement [28]. If the gap is longer, it might be motivated by factors that the researcher cannot control. However, if it is shorter, the gap may not give adequate time for adjustment in the perception process of individuals and their use of a new system [28].

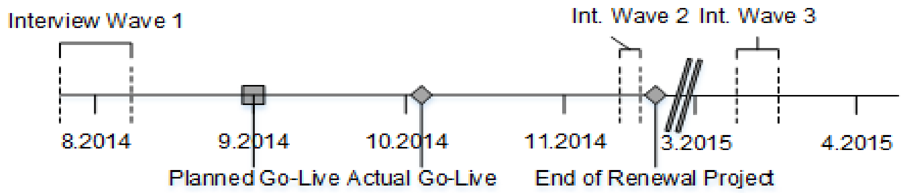


Figure 1. Data collection plan at FASHION

The first interview series started with an assessment of the system's version at the time and with an assessment of the typical adaptation of users with regard to the system. Furthermore, we asked for users' and managements' expectations regarding the introduction of a new PIMS release. In the second wave, we interviewed a content manager, the team leaders, a manager, and the technicians to assess their evaluation of the project and the actual progress made. This second round of six interviews, included questions whether the expectations were met by the new release. It also included questions about the user adaptation and the adaptation process necessary to deal with the new system shortly after its introduction.

Only the deputy department head could be interviewed during the mid-term sessions because of the ramifications of the busy holiday season. A content manager is the second missing interviewee interviewed before the project, as he had voluntarily dropped out of the company in the meantime. The third round of in total eight interviews included a final round of questions whether the expectations were met and questions regarding the adaptation. We asked users about the amount and kind of organizational support that they received in each round (see Table 1).

Table 1. List of Interviewees at FASHION

Interviewees	*Total # of Persons*	*Interviews Wave 1*	*Interviews Wave 2*	*Interviews Wave 3*
Content Managers	3	2	1	2
Team Leaders	2	2	2	2
Lead Technician	1	1	1	1
Deputy Technician	1	1	1	1
Deputy Department Head	1	1	1	1
Department Head	1	1	-	1
Totals	**9**	**8**	**6**	**8**

When possible, we used *participant observation* to gain a practical understanding of their interactions with the software. This aspect was supported by the previous role of the main researcher on this project, who had been an intern in the e-commerce department as a student. *Document analyses* were mainly part of the initial analysis during and after the first round of interviews and helped to understand the organizational structure, IT infrastructure, and IT architecture. In addition, we documented the rules and procedures of data collection in a case protocol to ensure rigor in data collection. Furthermore, a case study data base was used which contained the interview transcripts, field notes, collected documents, coded data, and the coding scheme [24, 29]. The data was organized based on Spradley's suggestions [30]. This organization allows separating objective facts in the condensed and expanded account and subjective interpretations in the analysis account and fieldwork journal.

3.3 Data Analysis

The unit of analysis of the case study is the work system of FASHION. For coding and tracking the qualitative data from the field, we used AtlasTI and followed an inductive coding approach. Inductive coding is appropriate in our research context as it allows to abstract themes, which are mentioned by interviewees on a reoccurring basis. We started with open coding of the interview transcripts. These open codes are descriptive and merely allow a categorization of constructs identified in the interview transcripts. We intensively compared and contrasted the developed categories with each other. In a second phase, we conducted axial coding to refine the interpretation of the categories and properties. At this stage, we also controlled for a possible researchers' bias in the categorization process by crosschecking the categorizations of the codes with an independent student assistant's categorization of a sample of three examples for each category. The categorization was very similar.

We used a critical-realism (CR) as the epistemological perspective for the analysis of our gathered data. CR distinguishes between a transitive and intransitive domain. The intransitive domain consists of the elements such as events and the causal powers in the ontological domain of the actual and the real that the researcher attempts to understand [31]. The transitive domain contains the observations, knowledge or theories about the independent world of the intransitive domain. A perfect match between theories and reality is not likely, and theories are fallible. Intransitive elements do not change over time, however, the theories about them do and presumably become less fallible [9]. Thus, this research approach is ideal for the analysis for complex interactions and consecutive smaller events [9], such as the forming of perceptions of project success in iterative steps.

Specifically, we followed the principles for conducting the CR research in IS by Wynn and Williams [9] for our data analysis. The first principle is the detailed explication of events through the abstraction of individual's experiences, as the foundation of causal analysis. This step is crucial for understanding the PIMS and FASHION as an organization. Second, we explained the structure and the contexts of these events. For instance, this involved the analysis of the sequence of the flow of information inside of FASHION. Third, in the process of retroduction we identified the hypothetical causal mechanisms, which could explain the specific occurrence of these events [9, 32]. Fourth, we evaluated with empirical corroboration whether the hypothesized mechanisms illustrate reality correctly, elucidate the events better than other mechanisms and are appropriate explanations with a high degree of causal power by referring to our obtained data [9]. This was executed by constantly referring back to our transcripts and the documents. Finally, we employed triangulation, mainly to emphasizes the necessity to use more than one source of evidence, that is in our case the combination of different interviewees insights with our observations and document analysis to find an appropriate causal explanation for the different perceptions of reality [9]. The resultant contributions of a CR study can be classified as type II theory [10], which provides explanations for the occurrence of a phenomenon in a social system [9].

4 Findings

At first, we are going to outline characteristics of the interviewees and the stakeholder groups relevant to the case study. All people interviewed knew PIMS since its launch in the firm in August 2013 and had several years of experience inside the firm. The content managers were working in a rather small team that was the innovative "new" group in the company. Management had been familiar with the PIMS since its introduction and had been in the business area for two to three years. The in-house technicians were relatively new to the area and the technology. The head technician had joined the firm half a year before the migration project, while the deputy technician joined just one month before the start of the renewal project. As head technician, he was responsible for planning and organizing the renewal project. However, his background had been more in e-shop-systems. There was a multitude of ongoing projects at the same time inside FASHION. His only in-house technical support was the new deputy technician, who had worked with the PIMS at a previous employer. But he had to familiarize himself with the renewal project and its scale.

Initially, the new release of the PIMS had been purchased. However, it had not been ready for the initial implementation. Management had made the decision to implement the old release with some upgrades, which were ready at the time, and to create a hybrid version. The organization was still incorporating that change, as the head technician noted*: "Just recently, we were at a user group meeting of the software producer. Based on their project status classification, we just finished the renewal phase and are currently entering stabilization. However, the new release will* *disrupt that phase."* The project team for the introduction of the new release included the deputy department head, the head technician, and the deputy technician. A contract for the new release was signed in April 2014, which also included the move to a new service partner. The deputy department head gathered 23 end-user requirements, i.e. their expectations for the release of PIMS's Web-Client version in meetings in April and May 2014. Table 2 presents the four main requirements of this list as assessed by the head technician after the end of the project. These four common themes of expected changes emerged during the first round of interviews with content managers: First, the users expected an adaptation of the user interface for the product classification process. This included a change from a slow drag and drop process of individual product classifications, up to 20 at a time, to a simultaneous selection out of a list of characteristics. Second, an improved semantic search was required for the Web-client. Third, seamless navigation between product, variant, and article level in the PIMS Web-client also featured in the interviews. Fourth, product images should be available on all presentation layers in the system. The department head had the following expectations: *"[Whenever the new release is migrated and running], we will start by introducing a new design of the content management process. This will be a project of another five days [...]."* This process has not been implemented to date, October 2016. Management's expectations were in clear divergence to the technicians' expectations. Both technicians mainly expected benefits for handling of the technology and background changes. Besides, they planned a 1:1 migration to the new service provider.

Table 2. Main requirements for PIMS assessed after project completion

Main Requirements	*Status at Project's End*
Seamless navigation of system levels (product, variant, article)	Done
Integration of a spelling check	Testing
Automatic classification of products	To Do
Product images visible on all presentation layers in the system	In evaluation

Subsequently, the two technicians and the renewal partner prepared technical changes for the actual project. The migration of the entire data for the PIMS to the new hosting and general service provider was planned for the end of August 2014 and the planned go-live was on the 1st of September 2014. Separate hosting and service partners characterized the previous set-up. A renewal of the new release involved a service partner, who hosts and provides maintenance services out of one hand. An attempt to go-live on the 1st of September was made. Soon after this go-live, users from the content management team experienced such a lack of system performance and data quality that the attempt was abandoned. The new release was deemed not ready and FASHION reverted to the old set-up for the rest of the month. A new attempt to go-live was made at the beginning of October. Even at this point, users soon experienced a severe lack of performance and responsiveness. The service partner had underestimated the server capacity necessary to run the old PIMS implementation. In-house technicians began to learn that the original data model was incorrect for the design of the standard software. It lacked stability and had a slower performance than planned as it was used as a calculation tool for stock levels and other data, which was against the original design brief of the standard software. The head technician summed up the situation in the following way: *„The guys [from the first implementation partner] just screwed up a little. They were not capable of implementing a PIMS, at least of this scale."* Despite many separate efforts by the new service partner and the in-house technicians, a lack of performance persisted. Several data exports and imports were redesigned to reduce the workload for the PIMS. Ultimately, the stability of the system improved with a sequence of hot-fixes and bug fixes that were issued by the software producer and the service partner. Thus, it was possible to overcome the worst part of stability and performance issues within the first two weeks of the new release. The deputy technician stated: *"At the moment, we are happy that the system runs in an identical version on the new platform."*

It is evident, that end-user expectations were not confirmed positively. When asked about the share of expectations that were met, a content manager stated: *„About 30 to 35%. [...]. Expected was 60% of fulfillment of requirements."* Hence, she subsumed: *"The product has improved a little."* The newly assigned team leader for content management commented: *"Currently, I would say that performance-wise we are back on the level of the old version of the PIMS."* Furthermore, it was noted that *"the new classification approach with drop-down lists takes longer now."* This was due to a lack of performance of the hardware with the new hosting partner. Thus, a goal of the renewal was initially missed. While management acknowledged these problems, the deputy department head had the following impression: *"You can feel it, they [the users] are also satisfied. Some of the things that have changed are things that they wanted. […] 30% of requirements were ready with the first version after the release. We are currently implementing another 20% of our requirements and the other 50% are extra goodies. They will follow later."*

A great variety of perceptions persisted to the end of the project. For instance, one content manager stated that she felt like only 20% of the requirements were actually met. Independently, the interviewed content managers and team leads stated similar figures. In an interview with the head technician after completion, it turned out that he had never been aware of the list of requirements from the workshop. Just one of the 23 requirements on the list was met over the duration of the project. Eight requirements were classified as a planned "To Do" by the lead technician, two more were being worked on or planned, while the rest of twelve requirements was not understood or seen as conflicting by him. A content manager subsumed that "*the performance after the introduction of the new release and the management of the transition issues is just back to the way it used to be with the old release […]*."

This perception of the overall progress did not square with management's perception. The managers felt that employees focused too much on the negatives and the deputy department head stated: *"The new release is still about 20 to 30% slower than the old release. Many employees focus on this downside during conversations."* His perception of the fulfillment of the initially gathered requirements was fundamentally different and more positive: *"I would presume that 50 to 60% of the requirements on our list have been met by now."* The attitude towards problems, which were raised by content managers, was clear: *"'Yes, everything was better in the past'. Yes, the change was not easy, it has brought additional workload, it also brought certain restrictions, but it was just necessary, […]."* The department head expressed a new idea of the initial project's focus: *"We are closer to the standard. We have almost 100% of the standard. This was the top priority."* This was a fundamental change to the beginning of renewal project when users were asked to formulate a list of requirements during the workshop. This raised user expectations, which were slowly crushed as the project progressed. The department head was aware of this, but did not inform the content managers or purchasing department end-users: *"[…] we did not ask for intensive feedback, because we implemented very little from the long list of requirements because we changed a lot in the backend instead. [...], we can invest more in features and usability [when the backend of the system is stable]."* This decision was made because the department head had been aware of the issues during the migration process: *"[…] after the introduction of the new release, we had catastrophic system performance."* Nevertheless, the department head was of the impression that individual performance had increased substantively: *"We have an increase of 30% in productivity and speed compared to the previous release."* Considering everything, he specified: *"I am convinced that we have a "Ferrari" [i.e., PIMS] that we cannot use appropriately."* All the while also stating: *"That is just not a perfect system and we probably expect too much of it."*

Yet, technicians' perception of the overall project was different at the end of the initial renewal phase. The deputy technician acknowledged: *"We carry a huge load of requirements that were not met previously. There are plans, but neither the time nor the ability to create To-Do's to actually assess and implement the desired changes."* However, some parts of the transition were also perceived to be a success as it was possible to reduce the time needed for data imports and exports for the PIMS by about 50%. In part, this can also be attributed to changes in the hardware of the hosting partner. The technician was of the impression that this new speed in reaction motivated some users: *"The system pretends to help me, that's cool."*

During the course of the renewal project, many new projects and issues had overridden what the technicians wanted to achieve regarding their preparation of technology. The head technician's description of his interaction regarding the management of IT projects with FASHION's management illustrates the discrepancy in thinking: *"Management certainly listens, however it is unclear whether they truly understand and take note when necessary. Our department head admitted to me that we probably addressed too many issues at once. Overload will lead to failures. […]".* When the head technician perceived these issues during the migration process, he made the momentous decision to scale down the project. The project complexity was reduced by focusing on the main migration: "*In the end, everything had to be rushed because our management had communicated a deadline inside our organization. […] It is my belief that internal policy created an expectation inside the organization which resulted in pressure and eventually lead to friction losses."* End-users were not clearly informed, which allowed the different expectations to linger.

5 Discussion

We identified several mechanisms that can help us to explain the evolving perceptions of end-users. An important mechanism, which influenced end-users' perceptions was the **narrative of success**, the focus in management's communication on motivating employees by showing them that they are successful and taking part in something meaningful for FASHION and the multi-channel retailer as a whole. All interviewed content managers perceived the new release as an opportunity to improve their work and its outcomes, and fundamentally believed in this narrative of success. As strong e-commerce growth required many rapid changes, they had developed a common culture of trial and error, which they all ascribed to themselves. This mechanism is closely linked to the insight that success and particularly project success is socially constructed and perceived by different stakeholders [5, 8] and can be linked to organizational culture [33]. The specific aspect of narratives of success has also been raised in previous research on IS projects [13, 34].

Furthermore, employees in FASHION's e-commerce department show a great deal of synergy and loyalty to each other, as well as to the department head. These are the aforementioned characteristics of a good team [21, 22]. However, in this situation, the mechanism of **hierarchical groupthink** was present on the basis of belief in the e-commerce department's narrative of success. The following quote of the department head illustrates his power in setting an agenda: *"We have spent the last three quarters with very intense discussions and got a lot of scolding: Everything was better before [with the system before PIMS]. I have heard [this] so often, but all have to agree to it or have to engage with it, because there is no alternative. Now everyone agrees with it."* Janis [35] provided six criteria to identify and determine a situation of groupthink: 1.) *Little or no consideration of alternate plans:* Management at FASHION did not have a back-up plan for a failed migration or further technical issues. For instance, downtimes were seen as a given. 2.) *Risk is not assessed:* Management and Technicians at FASHION did not assess the risk for the operations of the difficult migration that they planned. Subsequently, the migration and go-live of the new release failed. If people raised issues, it was stated that the project simply "had to be done in this way". 3.) *No review is taken of rejected plans:* There was just one option: The execution of the initial plan. This was further enforced by commercial arrangements for the release change, which had been designed by management inflexibly to save money. The failure of the first renewal attempt for the new release occurred, because the software of the new release had not been ready. 4.) *Advice from outsiders is not sought*: Management did not feel able to fund a specification project by technical experts from a consultancy. *5.) Facts that support the plan are acknowledged, facts that do not support the plan are ignored:* This was observed in management's attitude to end-users input regarding project success. From management's point of view end-users simply focused too much on the negative. 6.) *Contingency plans are not created:* There was no alternative plan created for the renewal project and the implemented solution. The technology is a *'Ferrari'* and simply not used properly. The described groupthink had the effect that content managers bought in to this assessment and that significantly influenced, how they made sense of the renewal project.

As aforementioned, there were four different levels of information: technicians, managers, content managers, and purchasing department end-users, who were not informed about the particularities of the project. These different groups had different sensemaking experiences. This is due to the different points in time at which they received their inputs. At first, the technicians became aware of the issues with the initial implementation. This was crucial for other parties' sensemaking. As the management was made aware of the technological issues, the deputy department head commented: *"If you turn one stone, you have to turn them all."* This meant scope creep and a more comprehensive change then initially anticipated, but also a change of priorities. As a consequence, the aforementioned list of requirements remained unknown to the lead in-house technician till one of the researchers presented it to him after the end of the project. The end-users realized a dawning failure based on the results they perceived in their daily work. A content manager commented: *"It became evident during the run of the project that our [the content management] team's wishes […] were difficult to implement."* The content management team reacted with not focusing on the wishes and expectations anymore. This can be described as the mechanism of **inherent fatalism** of end-users. Instead, they realized that the renewal project was a threat for their productivity. Overcoming the threat and the difficult phase was therefore a great success. According to a team lead, the new attitude to the project became: *"It simply had to be done."* She described their experience with the adversity as a *"state of war"*. She went on to say: *„It is a positive experience to go through such difficult periods. It is an opportunity to grow personally and to see what you are capable of."* We interpret the described personal growth and experience of performing against the odds as the seed for the perception of success that end-users reported. This appears to be at the heart of their sensemaking process. It overshadows the project and its original purpose over time. The other team lead stated in the third rounds of interviews: *"I do not know [how many requirements were met]. I have no idea. [...] You get used to situations and if something is suddenly missing from the tools that you use, you find other ways. […] Whenever you get used to something you stop questioning it. Hence, I do not know what can be improved at the moment."* The hallmark of success in such a scenario became reaching the previous level of performance and they abandoned the goal of renewal. As a group, the users at FASHION developed a reliable system, similar to those described in the literature [23], to cope with the adversity that they perceived because of the technological glitches in their work environment. Overcoming the situation as a group also gave them a collective mind and a collective feeling of success. This finding adds to previous research which has identified the importance of organizational culture for IS project success in general [33]. Aspects of inherent fatalism as a mindset, its antecedents, and its consequences have featured in previous research. For instance, research on perceived organizational support and psychological contracts of employees with their employer [36] has investigated conditions that might lead to inherent fatalism on the part of the employees. Part of the process to readily accept the situation in the workplace is the rationalization process of individual end-users. More specifically, motivated reasoning [37, 38], which is the reliance on a biased set of cognitive processes, is likely to be important for explaining end-users ability to focus on the aspects under their control. The end-users could have been motivated to avoid a reasoning that would stain the embraced narrative of success of FASHION. As a consequence, such an approach allows them to remain motivated to work [39] at FASHION. The organization relied on the described combination of mechanisms, which has its roots in the instilled organizational narrative of dynamism and success, to motivate users to overcome the problems in daily use. As a result, the deputy department head believed that all people involved were satisfied and summarized: *"The users found ways to deal with the performance problems."*

6 Conclusion

A CR approach enabled us to develop a better nascent theory for the understanding of various perceptions and evaluations of success of IS projects in organizations. Our explanation of the link between the mechanisms identified above is the main contribution of our study. We use them to explain the discrepancy between end-users' perception and real renewal project success: For *end-users*, the perceived success of overcoming the adversity of the renewal project was a good match with the overall groupthink, and the predominant organizational narrative. They perceived themselves as the group of people that was working in a dynamic market environment and as those who successfully struggle with its dynamism. Overall, their sensemaking of the situation had a fit with FASHION's organizational narrative. From this, we draw the conclusion that overcoming the adversity of a project's ramifications is a big factor in the perception of successful projects by end-users. This creates a feeling of unity and resolve in good teams. The greater purpose of being part of something interesting (a growing and dynamic business – fitting the organizational narrative) is also an important aspect. For *management*, the resilience of end-users, who are motivated in such a way, is crucial to ensure relative success to their adjusted objectives. As observed in our case, managers seem to adapt their level of perceived success based on the information they receive from the technicians, who are closest to the matter at hand, but are not necessarily aware about the overall story that has been told by management about the project they are working on. Thus, there is a wider disconnection in the sensemaking of individuals in an organization about the success of a project. As long as management dominates the perception of the business environment and end-users buy into the derived organizational narrative, it is likely to influence the sensemaking process of end-users. In our case, this means that the adversity of the initially *planned* technological change is seen as inevitable on the level of end-users. End-users seem to consider the greater cause inherent in the organizational narrative and respond with a fatalistic and resilient attitude and form a reliable system, which allows them to cope with the adversity related to technology project in their organization. For *technicians*, this means that their sensemaking is constrained by time pressure and in our case the inevitable lack of experience with the PIMS. In this situation, they had to make sense on the fly. Furthermore, they did not feel empowered to manage relationships with end-users and expectation management on their own. Overall, this led to the described situation in which the perception of the business environment and the resulting organizational narrative dominated the perception of a project's success. We think that this theoretical understanding is generalizable as the organizational narrative, which informs perception, is likely to depend on the organizational environment.

A possible limitation of a single case study is always generalizability. We deem a single case as appropriate for exploratory research and aim to challenge generalizability of our results on the basis of multiple cases in future research. It is a practical implication of this paper that managers should make sure that they actively nominate someone, who plays the role of a devil's advocate [22] to manage the expectations related to a synchronized plan. This will alleviate the problem of groupthink based on a similar perception of the environment and the resulting organizational narrative. In our particular case, the common believe led to a lowering of expectations which allowed to reinterpret failure as success in meeting adjusted expectations. This is a benevolent outcome. It is also possible, that the organizational narrative further aggravates end-users. A narrative told to motivate employees can ring hollow if it is not backed up by reality. Thus, management and technicians should communicate more directly and more transparently with end-users about the underlying technology. Even if they do not understand the technology in detail, they are likely to welcome the gesture of inclusion and the possibility to participate. In a different environment as in our case, users can resort to adverse behavior such as user resistance [40, 41]. The circumstances of this can be at the center of future research.

References

1. Manager-Magazin, http://www.manager-magazin.de/magazin/artikel/a-891269-3.html (Accessed: 02.11.2016)
2. Andersen, E.S.: Toward A Project Management Theory for Renewal Projects. Proj. Manag. J. 37, 15–31 (2006).
3. Atkinson, R.: Project Management: Cost, Time and Quality, Two Best Guesses and A Phenomenon, Its Time to Accept Other Success Criteria. Int. J. Proj. Manag. 17, 337–342 (1999).
4. Lech, P.: Time, Budget, And Functionality? - IT Project Success Criteria Revised. Inf. Syst. Manag. 30, 263–275 (2013).
5. Thomas, G., Fernández, W.: Success in IT Projects: A Matter of Definition? Int. J. Proj. Manag. 26, 733–742 (2008).
6. Basten, D., Joosten, D., Mellis, W.: Managers' Perceptions of Information System Project Success. J. Comput. Inf. Syst. 52, 12–21 (2011).
7. Bala, H., Venkatesh, V.: Changes in Employees' Job Characteristics During an Enterprise System Implementation: A Latent Growth Modeling Perspective. MIS Q. 37, 1113 - 1140 (2013).
8. Cecez-Kecmanovic, D., Kautz, K., Abrahall, R.: Reframing Success and Failure of Information Systems: A Performative Perspective. MIS Q. 38, 561–588 (2014).
9. Wynn, D., Williams, C.: Principles for Conducting Critical Realist Case Study Research in Information Systems. MIS Q. 36, 787–810 (2012).
10. Gregor, S.: The Nature of Theory in Information Systems. MIS Q. 30, 611–642 (2006).
11. Doherty, N.F., Ashurst, C., Peppard, J.: Factors Affecting the Successful Realisation of Benefits from Systems Development Projects: Findings from Three Case Studies. J. Inf. Technol. 27, 1–16 (2011).
12. Ewusimensah, K., Przasnyski, Z.H.: Factors Contributing to the Abandonment of Information-Systems-Development-Projects. J. Inf. Technol. 9, 185–201 (1994).
13. Fincham, R.: Narratives of Success and Failure in Systems Development, Brit. J. Mgmt. 13, 1–14, (2002).
14. de Wit, A.: Measurement of Project Success. Int. J. Proj. Manag. 6, 164–170 (1988).
15. Lech, P.: Time, Budget, And Functionality? - IT Project Success Criteria Revised. Inf. Syst. Manag. 30, 263–275 (2013).
16. Byrd, T.A., Thrasher, E.H., Lang, T., Davidson, N.W.: A Process-Oriented Perspective of IS Success: Examining the Impact of IS on Operational Cost. Omega. 34, 448–460 (2006).
17. Nicolaou, A.I., Masoner, M.M., Welker, R.B.: Intent to Enhance Information Systems as a Function of System Success. J. Inf. Syst. 9, 93–108 (1995).

18. Masoner, M.M., Lang, S.S., Melcher, A.J.: A Meta-Analysis of Information System Success: A Reconsideration of Its Dimensionality. Int. J. Account. Inf. Syst. 12, 136–141 (2011).
19. Lyytinen, K.: Expectation Failure Concept and Systems Analysts' View of Information System Failures: Results of an Exploratory Study. Inf. Manag. 14, 45–56 (1988).
20. Weick, K.E., Sutcliffe, K.M., Obstfeld, D.: Organizing and the Process of Sensemaking, Org. Sci. 16, 409-421 (2005).
21. Huczynski, A.A., Buchanan, D.A.: Organizational Behavior - An Introductory Text. Prentice-Hall, London, UK (1991).
22. McAvoy, J., Butler, T.: The Role of Project Management in Ineffective Decision Making Within Agile Software Development Projects. Eur. J. Inf. Syst. 18, 372–383 (2009).
23. Weick, K.E., Roberts, K.H.: Collective Mind in Organizations: Heedful Interrelating on Flight Decks. Adm. Sci. Q. 38, 357–381 (1993).
24. Yin, R.K.: Case Study Research: Design and Methods. (2009).
25. Myers, M.D.: Qualitative Research in Business & Management. Sage Publications, London, UK (2009).
26. Lawrence, M., Low, G.: Exploring Individual User Satisfaction within User-Led Development. MIS Q. 17, 195–208 (1993).
27. Myers, M.D., Newman, M.: The Qualitative Interview in IS Research: Examining the Craft. Inf. Organ. 17, 2–26 (2007).
28. Bhattacherjee, A., Lin, C.-P.: A Unified Model of IT Continuance: Three Complementary Perspectives and Crossover Effects. Eur. J. Inf. Syst. 24, 1–10 (2014).
29. Dubé, L., Paré, G.: Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations. MIS Q. 27, 597–636 (2003).
30. Spradley, J.P.: Participant Observation. Hol, Rinchar and Winston, New York (1980).
31. Zachariadis, M., Scott, S., Barrett, M.: Methodological Implications of Critical Realism for Mixed-Methods Research. MIS Q. 37, 855–879 (2013).
32. Mingers, J.: Realizing Information Systems: Critical Realism as an Underpinning Philosophy for Information Systems. Inf. Organ. 14, 87–103 (2004).
33. Jackson, S.: Organizational Culture and Information Systems Adoption: A Three-Perspective Approach. Inf. Organ. 21, 57–83 (2011).
34. Brown, A.D., Jones, M.R.: Doomed to Failure: Narratives of Inevitability and Conspiracy in a Failed IS Project. Organ. Stud. 19, 73–88 (1998).
35. Janis, I.L.: Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes. Houghton and Mifflin, Oxford, England (1972).
36. Aselage, J., Eisenberger, R.: Perceived Organizational Support and Psychological Contracts: A Theoretical Integration. J. Organ. Behav. 24, 491–509 (2003).
37. Kunda, Z.: The Case for Motivated Reasoning. Psychol. Bull. 108, 480–498 (1990).
38. Rousseau, D.M., Tijoriwala, S. a.: What's a Good Reason to Change? Motivated Reasoning and Social Accounts in Promoting Organizational Change. J. Appl. Psychol. 84, 514–528 (1999).
39. Gagne, M., Deci, E.L.: Self-Determination Theory and Work Motivation. J. Organ. Behav. 26, 331–362 (2005).
40. Kim, H.-W., Kankanhalli, A.: Investigating User Resistance to Implementation: a Status Quo Bias Perspective. MIS Q. 33, 567–582 (2009).
41. Rivard, S., Lapointe, L.: Information Technology Implementers' Responses to User Resistance: Nature and Effects. MIS Q. 36, 897–920 (2012).

What Teams Need to Be Clear about – an Activity Theoretical Perspective on Shared Understanding in Health IS Implementation

Andy Weeger[1], Annkatrin Ott-Schwenk[2]

[1] Neu-Ulm University of Applied Sciences, Neu-Ulm, Germany
andy.weeger@hs-neu-ulm.de
[2] University of Applied Sciences Ravensburg-Weingarten, Weingarten, Germany
annkatrin.ott-schwenk@hs-weingarten.de

Abstract. Shared understanding facilitates the implementation of IS and can help to prevent unintended consequences. However, research is hitherto not precise on the kind of knowledge such understanding needs to capture. Taking an activity theory perspective, this paper theorizes that shared understanding needs to cover knowledge, experiences and perspectives on the contested activity systems the HIS is implemented in. Analyzing the data of an in-depth case study, it is found that issues emerging during the rollout can be traced back to a lack of shared understanding about the affected activity systems, particularly to insufficient shared understanding about the instrument-mediated relationships between contradictory motives, rules and the evolved division of labor. These findings are synthesized in a framework on critical aspects of shared understanding. This framework offers a coherent explanation for the rise of unintended consequences and enhances our learning of shared understanding in IS implementation.

Keywords: Health IS, Implementation, Shared Understanding, Activity Theory

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Weeger, A.; Ott-Schwenk, A. (2017): What Teams Need to Be Clear about - an Activity Theoretical Perspective on Shared Understanding in Health IS Implementation, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 544-558

1 Introduction

Society expects a great deal from health information technology (IT), particularly concerning quality, safety, and efficiency of healthcare [1]. Imagine a hospital implementing a health information system (HIS) to fulfill these expectations. Its management follows the advice of text books and clarifies the problems the IS should tackle, builds consensus, considers various options, chooses an IS that meets its needs, plans implementation carefully, continually involves stakeholders and maintains strong commitment [2, 3]. Nonetheless, after the rollout, unintended consequences such as workflow challenges and resistance arise leading the intended outcomes untapped.

The literature reveals that this scenario is not unique, but that HIS implementation regularly disappoints users and stakeholders [2]. Implementing information systems (IS) that meet the difficulties of the healthcare domain and its various stakeholders is thus a major challenge [4], particularly since HIS must account for numerous social and technical factors that originate in a complex and evolving environment [2, 5]. In this regard, shared understanding enables interdisciplinary teams to implement IS effectively [6, 7]. Such shared understanding involves the recognition of diverse perspectives, resolution of differences in meaning and an agreement on the socio-technical requirements and the role of the IS within the activities it is employed [8].

However, research has to date not precisely worked out the kind of knowledge and experiences that need to be covered by shared understanding, which enables teams to efficiently implement IS and to avoid unintended consequences [8]. To guide design and implementation of HIS that are better able to meet expectations and thus improve HIS implementation outcomes, it is valuable to gain deeper understanding of the aspects of healthcare activities that shared understanding needs to cover. Consequently, this paper aims to contribute to our knowledge by exposing and classifying socio-cultural characteristics of healthcare activities and the various perspectives of stakeholders that need to be recognized and aligned during HIS implementation. Thus, we put forth following research question: *What are crucial aspects of healthcare activities that need to be covered by shared understanding during HIS implementation to avoid unintended consequences?*

To approach this question, we conducted an interpretative case study of an HIS implementation in a large German hospital. Grounded in Activity Theory (AT), we developed initial assumptions about aspects of healthcare activities that need to be recognized, shared, and considered during IS implementation. AT informed our understanding of the elements and complexities of collective activities, the significance of diverse perspectives and the role of HIS as a mediating instrument. The assumptions derived from AT did not only guide data collection and analysis, they were also constantly refined during this process. By means of iteratively interpreting data and refining the theoretical concepts, we could develop a coherent explanation of the rise of unintended consequences that emerged during and after the rollout of the HIS. These explanations then informed the development of a framework, which points to crucial aspects that need to be covered by shared understanding to avoid such problems. Before we lay out our theoretical stance, we briefly point up the complexities of HIS implementation and introduce the concept of shared understanding.

2 Theoretical Foundation

Stability and reliability are pivotal in healthcare settings [9]. IS implementation should thus carefully manage change to healthcare activities, which are increasingly complex, contested, highly hierarchical, and multidisciplinary in nature [9-11]. Concerning this matter, prior research identified that embedding generic HIS as "one size fits it all" is reasonably difficult [12, 13]. Rather, the design and functionalities of HIS and evolved practices need to be wisely adapted [10, 14], such that the HIS fits with the socio-organizational context. These adjustments need to span the complete implementation process [15]. Consequently, several literature reviews have identified management and user involvement as critical measures for the design and implementation of HIS [2, 16].

Users, managers and IT professionals need to be aware of individual knowledge that is crucial for the design and implementation of HIS such as knowledge on the capabilities and limitations of the HIS as well as perspectives regarding necessary adjustments to work practices [16]. Aligning HIS and established work practices successfully is thus contingent on shared understanding.

2.1 Shared Understanding

Shared understanding is commonly defined as "the overlap of understanding and concepts among group members" [17] and refers to the extent to which a basic understanding of a common subject matter exists [6]. The common subject matter in IS implementation is the creation and adaptation of an IS that is well aligned with existing and emerging practices [6, 18-20]. As our research focuses on HIS implementation, we define shared understanding here as overlapping mental representations of knowledge and experience that allows people involved in HIS implementation to form joint explanations and expectations of the HIS and how it affects healthcare activities.

It is assumed that an understanding of reality is primarily constructed in the mind of the individual by organizing and combining new experiences with existing experiences and knowledge [21]. Thus, there are multiple constructions of reality and "there is no objectively, right understanding on a certain object of interest that matches reality, but rather different conceptualizations that may 'fit' reality better or worse" [8]. However, research on shared understanding clearly shows that combining and aligning individual understandings of IT professionals and their stakeholders is important for successful IS implementation projects [see e.g. 6, 18, 19, 20, 22-24].

Shared understanding does not imply that people involved in IS implementation simply accumulate individual conceptualizations of reality. Rather, they need to share their perspectives, negotiate meanings, and agree on a mental representation they want to follow [8]. Moreover, it is assumed that teams hold several mental representations, which are usually framed as task- and team-related models and either cover knowledge on the task, the way the task is approached or the team [8, 25]. Besides these rather broad categories, IS research has hitherto barely worked out the detailed kind of task-related knowledge and experiences that are crucial for effective IS implementation [8].

Though prior research reveals that insufficient fit between the HIS and the socio-organizational context contributes to HIS implementation failures [14], the kind of knowledge about healthcare activities that needs to be enacted remains unclear. AT provides concepts to analyze collective IS-mediated activities and to approach this gap. AT exposes the interrelations of the constituents of work activities and thus illuminates the socio-organizational context. Hence, AT should help us to understand what needs to be recognized, shared, and considered during IS implementation, particularly concerning the role of the HIS within these activities. Below we offer a brief introduction to AT.

2.2 Activity Theory

Although AT is still rather unfamiliar in IS research, it has emerged as an important theory for understanding change of IS-mediated work activity [26]. AT relates the different conceptions of human activities and the material, mental, and social resources through which they are enacted [27, 28]. To frame these relations, AT introduces the activity system as analytical unit that covers how diverse actors work together [28].

The triangular activity system comprises the mutual aim of the activity (the *object*), all the people who are directed towards the object (the *subjects*), cognitive and materialized *instruments* used in activity to realize the outcome more efficiently, explicit and implicit *rules* that govern the work, the way tasks are distributed (the *division of labor*) and the wider *community* of practitioners that revolve and evolve around the object [29-32]. Instruments, rules, and the division of labor empower actors with experience and skills collected in the past, relate the subjects to the community and determine the possibilities and boundaries of their actions [21, 29, 30, 33].

The object as a key concept of AT refers to a physical or cognitive entity that is under construction, moving from a ‘problem space’ to a result or an outcome [28, 31]. As the “true carrier of motives of the collective activity systems”, the object takes shape and acquires its value by being transformed by multiple members of an activity system. Thus, the object is an enduring purpose of the activity and determines individual goals and actions through which, in turn, it may be achieved [31].

Though object-relatedness is a key characteristic of human activity [34], it is not free of contradictions. People frequently perceive difficulties in constructing a connection between the goals of their individual actions and the motive of the collective activity [34]. These problems stem from the multi-voided and contradictory nature of human activities [31]. AT views activity systems as an accumulation of multiple perspectives, traditions and interests, where the division of labor creates different positions of the participants and the instruments, rules, and conventions carry multiple layers and strands of history [35]. Such contradictions are exemplified by diverse perspectives on the patient who may be considered as a person to be helped or as a source of revenue. Likewise, HIS as one of multiple instruments employed in healthcare activities can be viewed as a resource to provide better healthcare or to generate higher revenues.

Such contradictions within activity systems cannot be observed directly, they can only be identified through their manifestations [36] such as tensions, disturbances, and breakdowns that destabilize activities and reveal inefficiencies [31, 37]. Moreover, contradictions are also viewed as driving change. Building on these ideas, we conceptualize IS implementation as a process of reconfiguring activity systems to resolve contradictions or mitigate tensions. IS implementation projects are thus only successful, when tensions, disturbances, and breakdowns could have been reduced.

Modifying elements in activity systems also bears the risk of new or amplified contradictions that may surface as unintended consequences. To avoid the rise of unintended consequences, these contradictions need to be identified, considered and proactively approached by means of adequate strategies and measures such as purposeful adaptations to the IS and other cultural historical elements of the activity.

2.3 Summary and Initial Assumptions

Literature shows that effective design and implementation of IS is contingent on shared understanding [see e.g. 6, 19, 20, 23, 24]. However, prior research is not precise on the kind of task-related knowledge, experiences, and perspectives shared understanding needs to cover [8]. Building on the concepts of AT, we assume that the design and implementation of IS not only requires knowledge of individual actions and interactions with the IS, but also of the joint activity the IS mediates. IT professionals, users and stakeholders thus need to share their individual knowledge and perspectives on the system of instruments employed in collective activities and rather invisible aspects such as rules and the division of labor as well as the strands of history they carry.

An understanding that combines divergent conceptions of the historically evolved and contested activity systems, their elements, and interrelations is expected to enable IT professionals and users to identify contradictions and to agree on the role of the HIS within these systems. Such shared understanding will enable them to anticipate and avoid unintended consequences during and after rollout. Vice versa, we expect that significant obstructions of work activities indicate that the people involved in IS design and implementation based their actions on insufficient shared understanding of the contested activity systems. Thus, they were not able to anticipate novel or amplified contradictions and resulting tensions as reflected by unintended consequences.

3 Methods

To identify the aspects of healthcare activities that need to be captured by shared understanding, we decided to take an interpretative perspective and to conduct a single, in-depth case study [38-40]. An interpretive lens acknowledges that people create their own subjective and inter-subjective meanings as they interact with the world [38]. Case studies are capable to provide an ample description and analysis of these perspectives [40]. Thus, the case study approach fits well with exploring crucial aspects of healthcare activities that need to be recognized, considered, and shared in this specific context.

In line with Eisenhardt's [41] recommendation for case study research, we framed our research question in the light of prior research. Moreover, we engaged in an iterative process of considering theoretical concepts from extant literature, developing assumptions about the characteristics of shared understandings, comparing the patterns identified with our theoretical deliberations and enhancing our theoretical perspective [39]. Below, we introduce the case, and describe data collection and analysis in detail.

3.1 Research Case

In 2011, a large teaching hospital in Germany with approx. 9,000 employees in 17 clinics and 40 specialist departments started an initiative to gradually implement an IS for care documentation (henceforth CareDoc) at almost all intensive-care stations such as the internal medicine, the gynecological clinic, and the psychiatric clinic. Replacing the former paper-based nursing documentation instrument, CareDoc was intended to enable the wards to comply with external requirements, such as increasing the efficiency of day-to-day clinical activities and optimizing reimbursement. From an AT perspective, the paper-based instrument should be replaced to resolve tensions between the abilities of the hitherto employed instruments and evolving aspects of the object of healthcare activities, particularly efficient provision of healthcare services.

The paper-based care documentation is ward-specific and essentially a large piece of paper. This instrument is used to document basic patient data such as demographics and vital parameters, medications, treatment, and nursing reports. CareDoc replicates and extends these functions as it, for instance, also enables users to enter orders and to manage schedules. Moreover, CareDoc provides interfaces to the hospital's central information system used to manage master patient data and accounting processes.

Responsible for most medical documentation (i.e., data entry), the nurses at the ward are the key-users of CareDoc and thus particularly relevant during requirements-elicitation. This fact is also reflected by the project team responsible for the adaptation and implementation of CareDoc at the wards. The team was led by an IT project manager and supported by two IT professionals and three nurses. The nurses were relieved of their day-to-day activities at different wards to a varying degree (25-75 %).

Depending on the specialization of the clinic, the vendor of CareDoc provided several basic templates that reflect best practices from other hospitals. In the early phase of the project, the project team involved various users, whose task was to learn how the templates need to be aligned with standard processes at the clinics and their wards. The resulting customized templates reflect approx. 80% of the functionalities needed in all clinics. The remaining 20% were identified as ward-specific and were added as customization prior to the rollout at the wards. The customizations were predominantly informed by the nurses. To gather the requirements, a standard implementing procedure was adopted for every ward. First, the project team and the staff at the ward, particularly the nurses, jointly created a specification document that reflects ward-specific requirements. The project team then customized CareDoc accordingly. Two weeks before the rollout of CareDoc on the ward, the users were trained to use the adapted version of CareDoc. During the rollout phase, which lasted about 14 days, members of the project team helped ward staff to implement CareDoc within their day-to-day work and to fix emerging issues. The rollout strategy, particularly the selection of CareDoc and the basic template development, included key success factors noted in the literature: selection of adequate technology, senior leadership, and continuous consultation of key-users during all project phases [2, 3]. Thus, this case gives us the opportunity to identify aspects of healthcare activities that are not easily recognized by best-practice measures, but are critical for avoiding unintended consequences.

3.2 Data Collection and Analysis

Data was collected between 11/15 and 03/16 at the addictive disorders ward of the psychiatric clinic. The addictive disorders ward was chosen for two reasons. First, from 2012 to 2015, CareDoc was already rolled out at nine other intensive-care wards. Thus, the project team has gained significant experience and benchmarks concerning the adaptation and implementation of CareDoc. Moreover, they have already refined the templates so that they sufficiently cover most standard requirements. Second, the nurses at this ward were particularly engaged during the requirements elicitation and unanimously agreed on the general requirements so that little problems were expected.

We took the role of an 'outsider researcher' during ward-specific requirements-collection, customization, and rollout. We triangulated data sources (i.e., interviewed key-users and project team members, talked to and observed nurses and physicians, and reviewed documents) to safeguard reliability and credibility of the data [39, 42, 43].

Table 1. Formal Interviews (approx. 9.5 hours)

Interviewee	*Interviews (total length)*	*Documentation*
PT01 (Project Manager., IT professional)	2 (45 min.)	Notes/memorized report
PT02 (Project Team, IT professional)	3 (90 min.)	Notes/memorized report
PT03 (Project Team, nursing background)	1 (40 min.)	Verbatim report
PT04 (Project Team, nursing background)	2 (50 min.)	Notes/memorized report
PT05 (Project Team, nursing background)	2 (50 min.)	Notes/memorized report
NU01 (Nurse at the ward)	3 (110 min.)	Verbatim report
NU02 (Nurse at the ward)	3 (70 min.)	Verbatim report
M01 (Ward Mngr., line manager of nurses)	3 (100 min.)	Verbatim report

The formal interviews were set up as semi-structured conversations [40]. The interview guideline contained open questions about the interviewees' knowledge and experiences on the affected activities and the role of CareDoc within. The interviews covered all available project members as well as involved key-users as assigned by management (Table 1) and aimed at exposing the knowledge, experiences, and perspectives the key-users and project team members shared. During the interviews, we gathered data on their conceptions of the healthcare activities, knowledge, and experiences regarding conventions and tools that are related to care documentation and that govern actions and interactions between the people involved. The formal interviews were recorded and transcribed whenever possible. Some of the interviewees felt uncomfortable with the recordings. However, during these interviews a comprehensive number of notes was taken and a report based on these notes was prepared immediately after the interviews.

To identify emerging issues and to collect data on the tensions within the activity systems that may cause these issues, one researcher was on-site during rollout (5 days, 9 hours per day) and observed nurses' use of CareDoc, related outcomes and problems that significantly obstructed work activities. In addition, we conducted a vast number of informal conversations with nurses, physicians, and the members of the project team to gain understanding on the issues we have observed and on underlying tensions. These casual conversations and observations were instantly recorded in the case diary.

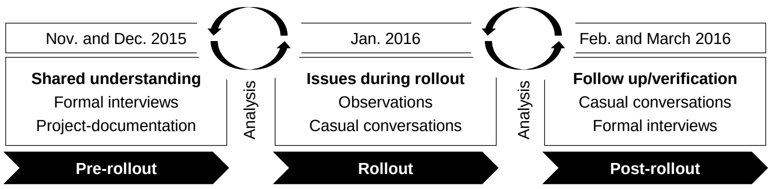


Figure 1. Iterative process of data collection and analysis

In accordance with the principles of interpretative research, data collection and data analysis occurred in a mode of continuous interplay, such that we were always open and willing to modify our initial assumptions (see Figure 1). One researcher constantly coded the data. The other researcher cross-checked the coding to ensure inter-rater reliability. Once agreement on the coding was reached, we compared, analyzed, interpreted and discussed emerging patterns as a team effort [42]. Whenever our interpretations differed, we went back to the data, theory, and/or the field and discussed the findings until we came up with an interpretation that was plausible for both authors [39]. Data collection lasted until all inconsistencies and gaps were resolved.

Aimed to understand which conceptions of the activities the IT project team and the nurses as key-users share, the first analysis cycle was conducted between the pre-rollout and the rollout phase. After coding the data on the individual conceptions of the work activity (i.e., open coding) [42], the resulting codes were analyzed and categories formed. The categories could have been assigned to an activity system component (e.g., subject or rule) or to a relationship between multiple components (e.g., relation between actors and implicit rules). Next, we compared the findings in each category to identify similarities, connections, and patterns between the individual conceptions.

During the second major cycle of data analysis, we focused on issues that emerged during the rollout and potential tensions between CareDoc and the healthcare activities. To do so, we coded data from the rollout-phase in search for aspects of the activity system that have previously not been shared as well as pointers to underlying tensions. Again, we created categorical codes, compared the resulting codes with the activity system components and assigned them to one of the activity system codes.

Finally, we analyzed the coded data and created a conceptual matrix [42]. For each problem that emerged, the matrix encompassed the aspects of the activity system that were covered by the shared understanding before the rollout as well as our AT-analysis of tensions within the activity system developed in the second cycle. This matrix allowed us to generate an overview of relevant data and to analyze how the emerging problems relate to aspects of the activities of nurses that were not covered by the shared understanding that guided adaptation and implementation of CareDoc.

4 Findings

The rollout strategy, including key success factors as noted in the literature [2, 3] helped to adapt CareDoc such that it enables nurses to document relevant data and to comply with standards. Moreover, CareDoc provides management and clinicians data better in quality (i.e., reimbursement and diagnosis). However, we also observed unintended consequences that emerged during and after rollout. These issues mitigated efficiency gains considerably and provoked resistance. Below, we analyze how these issues can be traced back to insufficient shared understanding of the activities affected by the HIS. We briefly present three of the most apparent unanticipated issues, illustrate the tensions and underlying contradictions, and analyze the aspects that were not covered by the shared understanding, which was developed during requirements elicitation and customization of CareDoc. Table 2 provides a summary of these findings.

Table 2. Summary of findings: contents of shared understanding (SU)

Aspects that were covered by SU	*Aspects that were not covered by SU*
Expected individual performance gains (e.g., documentation, reimbursement)	Contradictory perspectives (e.g., caring for patients vs. documentation quality and costs)
People/occupational groups involved in the activities that are affected by CareDoc	The historically evolved contradictory role of the community (e.g., volunteers and trainees)
The form and function of the formal paper-based documentation instruments and resulting technical requirements regarding types of data-fields, forms, reports, etc.	Additional, ward-specific enhancements of the formal instruments, their role as governance mechanism (e.g., reminder, overview), and the limitations of CareDoc (providing overview)
Legal requirements on healthcare activities (e.g., transparency, authentication, and medication handling)	Interpretation and evolved (contradictory) implementation of rules as reflected by norms, the instruments, and the division of labor

The first vignette refers to the coevolution of the instruments and other aspects of healthcare activities such as handover of tasks and performing medication in cases of emergency. The formal official paper-based documentation templates did not enable nurses to get a quick overview of all vital parameters, medication, and observations of the patients at the ward. To work around these shortcomings, over the years, nurses created and gradually refined 'monitoring sheets' for every patient at the ward. These letter-sized sheets hung at the board, enabled nurses to keep track of patients (e.g., in case of emergency and during shift handovers) and offered great flexibility. Data such as vital parameters, adapted dosages or important events were recorded on these sheets before typically trainees or voluntaries updated the formal documentation.

During requirements elicitation, it was decided to replace the monitoring sheets with CareDoc reports. Recording and analyzing data directly in CareDoc should reduce errors and enable early access to care documentation. However, the reports offered were not able to occupy the role of the monitoring sheets sufficiently: the report must be manually activated, is only accessible via the small monitors at the wards and does thus not offer a quick overview. For instance, NU02 stated that *"the monitoring sheets enabled us to overview patients much more easily […] which is quite important at a closed psychiatric ward."* Moreover, the reports did not occupy the function of a to-do list and reminder in the way as the monitoring sheets did. In consequence, shift handover was considerably obstructed and nurses' willingness to utilize CareDoc decreased.

Though the tensions between the adapted set of instruments and the nurses began to surface early, the project team, ward mgmt., and physicians were not able to understand why nurses requested to keep the monitoring sheets until the rollout. Particularly physicians, who presume timely and accurate data, advocated to use the CareDoc reports. This indicates that the project team did not thoroughly consider the role of the monitoring sheets as an important mediator between the actions and interactions of the nurses. Thus, these neglected the historically collected skills that enable nurses to deal with specifics at the wards (e.g., adapting medication, keeping track of patients).

The second vignette concerns medication management at the ward. Legally, nurses are only allowed to administer drugs as prescribed by a physician. Since physicians are frequently not available at the addictive disorders ward of the psychiatric clinic on short notice, however, nurses are in the need to adapt the dosage on their own and get authorization retrospectively (e.g., in case of an acute delirium or when symptoms like restlessness or shivering appear). This well-practiced shift in the division of labor has become an accepted norm at the wards. Since they feel less comfortable setting the dosages without consulting a nurse, physicians even asked if it *"is possible to calculate the optimal dosage within CareDoc*" (observation note).

Though, physicians, nurses, and IT professionals discussed the need for dynamic dosing intervals, customized related forms accordingly and agreed that the dosage could not be 'calculated', their shared understanding did not enable them to anticipate that CareDoc considerably restricts the flexibility as given by the paper-based instrument. These historically evolved aspects of the medication activity, particularly the tension between the way tasks are distributed in case of emergency and the legal requirements, were not uncovered during requirements engineering. Neither the nurses shared their knowledge and experiences regarding these aspects, nor did the IT project team point out that CareDoc will increase transparency due to time stamps of physicians' authorization of medication and nurses' administration of medication and thus, most likely, limits flexibility regarding medication handling at the ward.

Only during the rollout, the nurses recognized that CareDoc does not allow them to adapt the administration of medications on short notice or to enter up physicians' authorization after administering adapted medication without causing legal problems and –at worst– risking adverse personal consequences. Awareness of these problems negatively impacted nurses' attitude towards CareDoc. This even caused some of them to resist using the system until it was clarified how to deal with medication management in cases of emergency.

The third vignette reflects how gradual changes to the object of healthcare affected the division of labor. During the last years, cost pressures, and workload significantly increased at the wards, which stressed the need to gain efficiency. Therefore, the way tasks are distributed was continuously adapted. Amongst others, volunteers and trainees are increasingly asked to measure and document vital parameters. Legally, this must be supervised and authorized by a certified nurse, who also signs the documentation. In practice, however, nurses have not the time to accompany the volunteers and trainees and thus only sign the documentation afterwards.

Due to an IT policy, volunteers do not possess a user account that is required to perform documentation in CareDoc. Thus, after the rollout of the IS, documentation of vital parameters was limited to nurses and trainees. This obstructed the evolved mode of task distribution and increased nurses' documentation effort considerably. Since they either need to record vital parameter on their own or must sign recordings of trainees, they had less time to care for patients, which is usually the true motive of their activity (*"we spend too much time at the computer and have less time for the patients"*, NU01). Moreover, some of the nurses did not want to key in data they have not collected (*"I will not key in data that I have not measured"*, NU03).

Data collected prior to the rollout reveals that nurses and IT professionals did not share knowledge concerning the nature, history, and significance of the mode of task distribution at the wards and how CareDoc enforces conflicting legal requirements. During requirements elicitation, IT professionals and nurses did not consider how CareDoc could amplify tensions between legal requirements and the division of labor. Thus, neither IT professionals could anticipate that the characteristics of the IT tool would destabilize established ways of doing, nor ward management could clarify roles and responsibilities and thus take measures to mitigate or resolve these tensions.

5 Discussion

The paper-based instrument was found to limit the efficiency of healthcare activities (e.g., documentation quality, reimbursement). These tensions should be resolved by an HIS. Selection of an adequate system, senior leadership and user involvement enabled the project team to adapt the HIS so much that it covers most requirements, particularly those that concern the interactions of individuals with the system. However, as with similar HIS implementation initiatives reported in prior literature [2], the replacement of the legacy tools provoked unintended consequences. These were not anticipated, although the people involved could build upon benchmark data from nine rollouts prior to the addictive disorders ward and templates that are based on best practices from other hospitals. Case study data reveals that IT professionals and stakeholders failed to create shared understanding on critical aspects of the socio-technical environment the HIS is embedded in. The framework presented below summarizes and classifies these.

5.1 Crucial Aspects of Healthcare Activities

The unintended consequences we have observed indicate that users, managers, and IT professionals had difficulties to share –often– tacit knowledge and experiences related to three characteristics of their collective healthcare activities, particularly those that go beyond individual requirements. Accordingly, they struggled to anticipate how these aspects relate to the novel HIS, particularly how contradictions within the activity systems are mitigated or amplified by HIS implementation. In Figure 2 these aspects are classified (1-3) and located in a triangular activity system diagram (1-3).

The first aspect reflects that healthcare activities are characterized by contrasting objectives (aspect 1). Data reveals that the way the legacy tool and the IS account for different perspectives on the object of the collective activity was hardly considered (e.g., maintaining monitoring sheets vs. entering data directly in the HIS and using the reports). We found that exposing the contested nature of the activities and bearing in mind how the instruments account for divergent motives and thus how they impact the relationship between professional groups is critical. Creating shared understanding on this aspect would have most-likely enabled actors to anticipate and avoid the issues.

Second, patient care is characterized by a tension between the need for stability and sensitivity for variation [9]. This tension is also reflected in the historically evolved division of labor, which provides stability but needs to be sensitive to variation (e.g., in case of an emergency). Employed instruments need to account for these evolved and diverse modes of task distribution (aspect 2). Thus, actors must ensure that the HIS can govern routines and variations like their predecessors did or adapt the division of labor accordingly (e.g., task distribution between nurses, trainees, and volunteers).

Third, healthcare is highly influenced by regulation and traditional hierarchies [9]. However, evolved practices may contradict these conventions. Since HIS are often designed to comply with formal rules, they have the potential to amplify these contradictions. Hence, people engaged in IS implementation need to recognize, share, and consider explicit and implicit rules that govern actions and interactions between the subject and the community and clarify the role of the HIS within (aspect 3).

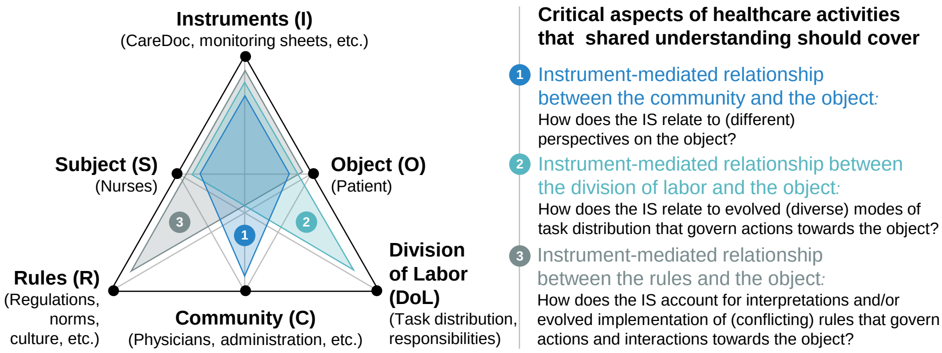


Figure 2. Aspects of healthcare activities that should be covered by shared understanding

The framework can help practitioners to more carefully consider the complex mediating relationships between instruments, rules, and the division of labor, facilitate the creation of shared understanding on the role of the HIS within and enables them to resolve contradictions without amplifying existing or creating novel ones.

Though awareness of the critical aspects will help, practitioners need to bear in mind that learning about an activity is at large inherent to the participation in an activity [44]. Thus, creating shared understanding might require IT professionals to participate in activities that are or may be affected by the HIS. This will most-likely enable them to better illuminate the aspects identified here, to anticipate emerging issues and to develop corrective actions. Practitioners might employ strategies such as cooperative prototyping, where users and IT professionals are involved [45]. Mutual prototyping may raise issues and shifts questions from subject-object interactions towards the rather invisible IS-mediated aspects of collective activities. Moreover, best-practices and approaches like the "MindMerger" [8] could benefit from methods that put an emphasis on the collective nature of human activity such as "expansive visibilization" does [31].

5.2 Limitations

Although this research offers a coherent explanation for the rise of unintended consequences after IS implementation in a complex setting, some limitations need to be acknowledged. First, the empirical data analyzed in this research stem from a single in-depth case study. Thus, it has to be shown if the activity theoretical perspective on shared understanding also explains outcomes in other settings. Second, data about early phases of the project could only be collected retrospectively, particularly regarding the design of the standard templates. Third, the timeframe for data collection was restricted. Thus, the rise of additional unintended consequences after data collection cannot be ruled out. Fourth, the number of interviews conducted is limited, some of the interviewees did not feel comfortable with recordings and only one researcher could monitor the rollout. Fifth, only the three most significant issues that emerged during and after rollout are reported in this paper. Though the analysis of the other outcomes support the conclusions drawn, the theoretical concepts still need to demonstrate their empirical validity [46]. Further research employing multiple and longitudinal case studies in other settings or quantitative studies that operationalize and test the theoretical statements may increase the confidence in the mechanisms identified here.

6 Conclusion

We know that effective design and implementation of HIS is contingent on shared understanding [see e.g. 6, 19, 20, 23, 24]. However, our learning on the kind of task-related knowledge and perspectives that need to be shared is limited. Our research contributes to this gap by providing a theoretically grounded framework that captures crucial aspects of healthcare activities that should be recognized, shared, and considered during HIS implementation projects. Moreover, we contribute to prior research by increasing our understanding on the socio-organizational aspects HIS must fit with [14]. We reveal that the people involved in HIS implementation need to create shared understanding about socio-organizational aspects that are not necessarily exposed using best-practice approaches. Following AT, the people involved need to recognize that HIS occupies positions within complex, collective, and historically evolving activity systems [35]. To resolve contradictions within these systems without provoking novel ones, the multiple perspectives and traditions that are carried by its actors, its artefacts, rules, and conventions need to be recognized, shared and aligned.

References

1. Jones, S.S., Rudin, R.S., Perry, T., Shekelle, P.G.: Health Information Technology: An Updated Systematic Review with a Focus on Meaningful Use. Annals of Internal Medicine 160, 48-54 (2014)
2. Cresswell, K., Sheikh, A.: Organizational Issues in the Implementation and Adoption of Health Information Technology Innovations: An Interpretative Review. Int J Med Inform 82, e73-86 (2013)

3. Cresswell, K.M., Bates, D.W., Sheikh, A.: Ten Key Considerations for the Successful Implementation and Adoption of Large-Scale Health Information Technology. Journal of the American Medical Informatics Association 20, e9-e13 (2013)
4. Kellermann, A.L., Jones, S.S.: What It Will Take to Achieve the as-yet-Unfulfilled Promises of Health Information Technology. Health Affairs 32, 63-68 (2013)
5. Kilsdonk, E., Peute, L.W., Riezebos, R.J., Kremer, L.C., Jaspers, M.W.: Uncovering Healthcare Practitioners' Information Processing Using the Think-Aloud Method: From Paper-Based Guideline to Clinical Decision Support System. Int. J. Med. Inform. 86, 10-19 (2016)
6. Tesch, D., Sobol, M.G., Klein, G., Jiang, J.J.: User and Developer Common Knowledge: Effect on the Success of IS Development Projects. Int. J. Proj. Manag. 27, 657-664 (2009)
7. Charaf, M.C., Rosenkranz, C., Holten, R.: The Emergence of Shared Understanding: Applying Functional Pragmatics to Study the Requirements Development Process. Information Systems Journal 23, 115-135 (2013)
8. Bittner, E.A.C., Leimeister, J.M.: Creating Shared Understanding in Heterogeneous Work Groups: Why It Matters and How to Achieve It. Journal of Management Information Systems 31, 111-144 (2014)
9. Fichman, R.G., Kohli, R., Krishnan, R.: Editorial Overview—the Role of Information Systems in Healthcare: Current Research and Future Trends. Information Systems Research 22, 419-428 (2011)
10. Goh, J.M., Gao, G., Agarwal, R.: Evolving Work Routines: Adaptive Routinization of Information Technology in Healthcare. Information Systems Research 22, 565-585 (2011)
11. Paul, R.J., Ezz, I., Kuljis, J.: Healthcare Information Systems: A Patient-User Perspective. Health Systems 1, 85-95 (2012)
12. Oborn, E., Barrett, M., Davidson, E.: Unity in Diversity: Electronic Patient Record Use in Multidisciplinary Practice. Information Systems Research 22, 547-564 (2011)
13. Poon, E.G., Blumenthal, D., Jaggi, T., Honour, M.M., Bates, D.W., Kaushal, R.: Overcoming Barriers to Adopting and Implementing Computerized Physician Order Entry Systems in U.S. Hospitals. Health Affairs 23, 184-190 (2004)
14. Ammenwerth, E., Iller, C., Mahler, C.: IT-Adoption and the Interaction of Task, Technology and Individuals: A Fit Framework and a Case Study. BMC Med Inform Decis Mak 6, 3 (2006)
15. Heeks, R.: Health Information Systems: Failure, Success and Improvisation. Int. J. Med. Inform. 75, 125-137 (2006)
16. Rahimi, B., Vimarlund, V., Timpka, T.: Health Information System Implementation: A Qualitative Meta-Analysis. Journal of Medical Systems 33, 359-368 (2008)
17. Mulder, I., Swaak, J., Kessels, J.: Assessing Group Learning and Shared Understanding in Technology-Mediated Interaction. Educational Technology & Society 5, 35-47 (2002)
18. Levesque, L.L., Wilson, J.M., Wholey, D.R.: Cognitive Divergence and Shared Mental Models in Software Development Project Teams. Journal of Organizational Behavior 22, 135-144 (2001)
19. He, J., Butler, B.S., King, W.R.: Team Cognition: Development and Evolution in Software Project Teams. Journal of Management Information Systems 24, 261-292 (2007)
20. Hoffmann, A., Bittner, E., Leimeister, J.: The Emergence of Mutual and Shared Understanding in the System Development Process. In: Doerr, J., Opdahl, A. (eds.) Requirements Engineering: Foundation for Software Quality, vol. 7830, pp. 174-189. Springer Berlin Heidelberg (2013)
21. Vygotsky, L.S.: Mind and Society: The Development of Higher Psychological Processes. Harvard University Press, Cambridge, MA (1978)
22. Schmidt, C.T., Kude, T., Heinzl, A., Mithas, S.: How Agile Practices Influence the Performance of Software Development Teams: The Role of Shared Mental Models and Backup. International Conference on Information Systems, Auckland (2014)

23. Faraj, S., Sproull, L.: Coordinating Expertise in Software Development Teams. Management Science 46, 1554-1568 (2000)
24. Hsu, J.S.-C., Chu, T.-H., Lin, T.-C., Lo, C.-F.: Coping Knowledge Boundaries between Information System and Business Disciplines: An Intellectual Capital Perspective. Information & Management 51, 283-295 (2014)
25. Cannon-Bowers, J.A., Salas, E., Converse, S.: Shared Mental Models in Expert Team Decision Making. In: Castellan, N.J. (ed.) Individual and Group Decision Making: Current Issues, pp. 221-246. Lawrence Erlbaum, Hillsdale, NJ (1993)
26. Karanasios, S., Allen, D., Finnegan, P.: Information Systems Journal Special Issue On: Activity Theory in Information Systems Research. Information Systems Journal 25, 309-313 (2015)
27. Blackler, F.: Knowledge, Knowledge Work and Organizations: An Overview and Interpretation. Organization Studies 16, 1021-1046 (1995)
28. Engeström, Y.: Expansive Learning at Work: Toward an Activity Theoretical Reconceptualization. Journal of Education and Work 14, 133-156 (2001)
29. Kaptelinin, V.: The Object of Activity: Making Sense of the Sense-Maker. Mind, Culture, and Activity 12, 4-18 (2005)
30. Nicolini, D., Mengis, J., Swan, J.: Understanding the Role of Objects in Cross-Disciplinary Collaboration. Organization Science 23, 612-629 (2012)
31. Engeström, Y.: Expansive Visibilization of Work: An Activity-Theoretical Perspective. Computer Supported Cooperative Work (CSCW) 8, 63-93 (1999)
32. Engeström, Y.: Learning by Expanding. An Activity-Theoretical Approach to Developmental Research. Orienta-Konsultit, Helsinki (1987)
33. Kuutti, K.: Activity Theory as a Potential Framework for Human-Computer Interaction Research. In: Bonnie, A.N. (ed.) Context and Consciousness, pp. 17-44. Massachusetts Institute of Technology (1995)
34. Leont'ev, A.N.: Activity, Consciousness, and Personality. Prentice-Hall, Englewood Cliffs, New Jersey (1978)
35. Engeström, Y., Punamäki, R.-L.: Perspectives on Activity Theory. Cambridge University Press, Cambridge (1999)
36. Engeström, Y., Sannino, A.: Discursive Manifestations of Contradictions in Organizational Change Efforts. Journal of Organizational Change Management 24, 368-387 (2011)
37. Allen, D.K., Brown, A., Karanasios, S., Norman, A.: How Should Technology-Mediated Organizational Change Be Explained? A Comparison of the Contributions of Critical Realism and Activity Theory. MIS Quarterly 37, 835-854 (2013)
38. Orlikowski, W.J., Baroudi, J.J.: Studying Information Technology in Organizations: Research Approaches and Assumptions. Information systems research 2, 1-28 (1991)
39. Walsham, G.: Interpretive Case Studies in IS Research: Nature and Method. European Journal of Information Systems 4, 74-81 (1995)
40. Yin, R.K.: Case Study Research: Design and Methods. SAGE, Thousand Oaks, CA (2009)
41. Eisenhardt, K.M.: Building Theories from Case Study Research. Academy of Management Review 14, 532-550 (1989)
42. Miles, M.B., Huberman, A.M.: Qualitative Data Analysis: An Expanded Sourcebook. Sage Publications, Thousands Oaks (1994)
43. Walsham, G.: Doing Interpretive Research. European Journal of Information Systems 15, 320-330 (2006)
44. Greig, G., Entwistle, V.A., Beech, N.: Addressing Complex Healthcare Problems in Diverse Settings: Insights from Activity Theory. Soc Sci Med 74, 305-312 (2012)
45. Mogensen, P.: Towards a Provotyping Approach in Systems Development. Scandinavian Journal of Information Systems 4, 31-53 (1992)
46. Lee, A.S., Baskerville, R.L.: Generalizing Generalizability in Information Systems Research. Information Systems Research 14, 221-243 (2003)

Zwischen Schattendasein, Governance und Entrepreneurship - Eine empirische Bestandsaufnahme zum Professionalisierungsgrad des IT-Managements in deutschen Krankenhäusern

Jan-David Liebe[1], Oliver Thomas[2], Franziska Jahn[3], Christian Kücherer[4], Moritz Esdar[1], Jan-Patrick Weiß[1], Jens Hüsers[1] und Ursula Hübner[1]

[1] Hochschule Osnabrück, Forschungsgruppe Informatik im Gesundheitswesen, Osnabrück, Deutschland
{j.liebe,m.esdar,j.p.weiss,j.huesers,u.huebner}@hs-osnabrueck.de
[2] Universität Osnabrück, Fachgebiet Informationsmanagement und Wirtschaftsinformatik, Osnabrück, Deutschland
{oliver.thomas}@uni-osnabrueck.de
[3] Universität Leipzig, Institut für Medizinische Informatik, Statistik und Epidemiologie, Leipzig, Deutschland
{franziska.jahn}@imise.uni-leipzig.de
[4] Universität Heidelberg, Institut für Informatik, Heidelberg, Deutschland
{christian.kuecherer}@informatik.uni-heidelberg.de

Abstract. Bei der Umsetzung der digitalen Transformation bewegt sich das IT-Management in Krankenhäusern in einem Spannungsfeld aus historisch-kulturellen Vorbedingungen und den besonderen Herausforderungen wissensintensiver Expertenorganisation. Um zu untersuchen, wie professionell das IT-Management vor diesem Hintergrund ist, wurde in der vorliegenden Studie der Professionalisierungsgrad des IT-Managements als Beschreibungsgröße vorgeschlagen. Darüber hinaus wurden Ausprägungen der IT-Governance und des IT-Entrepreneurships als mögliche Determinanten des Professionalisierungsgrades konzeptionalisiert. Ein entsprechend aufgestelltes, hypothesengeleitetes Untersuchungsmodell wurde anhand der Daten von 164 CIOs deutscher Krankenhäuser überprüft. Die Ergebnisse der Studie deuten auf Professionalisierungspotenziale des IT-Managements im strategischen und evaluierenden Bereich hin. Etablierte Kommunikationskanäle zwischen CIO und Krankenhausleitung sowie eine ausgewiesene IT-Budgetverantwortungen wirkten sich positiv auf den Professionalisierungsgrad aus. Zudem Das agierte das IT-Management umso professioneller, je stärker der IT-Entrepreneurship auf organisatorischer und individueller Ebene ausgeprägt war. Die Ergebnisse können den theoretischen Erkenntnisstand über die Wirkungsweise von IT-Governance und IT-Entrepreneurship erweitern und auf ähnliche, wissensintensive Expertenorganisationen übertragen werden.

Keywords: IT-Management, Krankenhaus, CIO, Governance, Entrepreneurship

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Liebe, J.-D.; Thomas, O.; Jahn, F.; Kücherer, C.; Esdar, M.; Weiß, J.-P.; Hüsers, J.; Hübner, U. (2017): Zwischen Schattendasein, Governance und Entrepreneurship - Eine empirische Bestandsaufnahme zum Professionalisierungsgrad des IT-Managements in deutschen Krankenhäusern, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 559-573

1 Einleitung

Die digitale Transformation birgt vielfältige Potenziale für eine Effizienz- und Qualitätssteigerung der Gesundheitsversorgung [1]. Zentrale Akteure innerhalb des Versorgungsgeschehens sind Krankenhäuser, die alleine im Jahr 2014 etwa neunzehn Milliarden Gesundheitsdienstleistungen erbracht haben [2]. Die Digitalisierung der hierfür notwendigen krankenhausinternen Abläufe soll die Planung und Koordination der Behandlung effizienter gestalten, die Informationstransparenz klinischer Prozesse erhöhen und den Zugang zu medizinischem Wissen vereinfachen [3]. Die Digitalisierungsmaßnahmen reichen dabei von der Dokumentation des Behandlungsverlaufs, über die Befundanforderung und -rückmeldung, bis hin zum Aufbau von entscheidungsunterstützenden Systemen, Wissensdatenbanken und der elektronischen Patientenakte (EPA) [4]. Bei der Durchführung entsprechender Maßnahmen bewegt sich das zuständige IT-Management der Krankenhäuser in einem Spannungsfeld aus historisch-kulturellen Vorbedingungen und gesetzlichen sowie ökonomischen Rahmenbedingungen.

Krankenhäuser sind, historisch bedingt, durch linienförmige Organisationsstrukturen und starre Hierarchien geprägt [5]. Abstimmungen über interdisziplinäre, abteilungsübergreifende Abläufe, welche ein primäres Betätigungsfeld des IT-Managements darstellen [6], finden zumeist nur auf oberster Hierarchieebene, zwischen der ärztlichen, pflegerischen und kaufmännischen Krankenhausleitung (KHL), statt [7]. Hinzu kommt, dass die vorherrschende berufsständische Autonomie, insbesondere des ärztlichen Bereichs, zu Zieldivergenzen und einer Vermischung von zentralen und dezentralen Entscheidungen führen kann [5]. Entsprechend können IT-Vorhaben in Krankenhäusern selten per Mandat durchgesetzt werden, sondern sind auf langfristige Abstimmungsprozesse angewiesen [8,9]. Vor diesem Hintergrund steht das IT-Management vor der Herausforderung, hochkomplexe und wissensbasierte Behandlungsprozesse in geeigneter Form informationstechnologisch zu unterstützen [10]. Hierbei wird der Erwartungsdruck durch den exponentiell steigenden, medizinischen Wissenszuwachs zusätzlich erhöht [11]. Bei der Integration entsprechender Anwendungen sieht sich das IT-Management nicht selten mit veralteten, heterogenen Systemen konfrontiert [12]. Eine weitere krankenhausspezifische Herausforderung ergibt sich aus den diversen gesetzlichen Vorgaben und Normen (bspw. Basel II, Medizinprodukte-Gesetz), welche die Etablierung eines umfassenden Risikomanagements erfordern [13] und somit in ihrer Umsetzung umfangreiche personelle und finanzielle Ressourcen binden können. Schließlich herrscht in der Krankenhauslandschaft ein zunehmender Wettbewerb und ökonomischer Druck [5], wodurch sich konkrete Erwartungen an das IT-Management ergeben. So sollen einerseits Effizienzreserven identifiziert und ausgeschöpft werden (z.B. die Vermeidung von Doppeluntersuchungen), andererseits steht die IT als potenzieller Kostenfaktor selber auf dem Prüfstand [12].

Die beschriebenen Rahmenbedingungen stellen das IT-Management der Krankenhäuser vor ein Dilemma: Zum einen erfordern die vielfältigen, teilweise divergierenden Herausforderungen weitreichende Handlungsfreiräume auf operativer, taktischer und strategischer Ebene [10], auf der anderen Seite können insbesondere historisch, kulturell und ökonomisch bedingte Restriktionen diese Freiräume begrenzen [5].

Über die Handlungsfähigkeit des IT-Managements von Krankenhäusern, im Sinne von planenden, steuernden und überwachenden Aktivitäten [10], existieren bis dato keine empirisch fundierten Erkenntnisse. Vielmehr wird dieser Forschungsgegenstand bisher nur randläufig in Studien zur IT-Adoption und zur IT-Governance betrachtet [1,4,9,14,15]. In der vorliegenden Arbeit wird zur Beschreibung des Umfangs und des Formalisierungsgrades des IT-Managements der *Professionalisierungsgrad* als konzeptionelle Beschreibungsgröße vorgeschlagen. Die Studie folgt weiterhin der Annahme, dass Ausprägungen der IT-Governance und des IT-Entrepreneurships mögliche Einflussgrößen auf den Professionalisierungsgrad darstellen. Entsprechend wurde vor dem Hintergrund des krankenhausspezifischen Kontextes ein hypothesengeleitetetes Untersuchungsmodell entwickelt, auf Basis dessen folgende Forschungsfragen beantwortet werden sollen: (1.) Wie ausgeprägt ist der Professionalisierungsgrad des IT-Managements in den deutschen Krankenhäusern und (2.) inwiefern beeinflusst die Ausprägung der IT-Governance und des IT-Entrepreneurships den Professionalisierungsgrad?

Durch die Betrachtung der Krankenhäuser, welche stellvertretend für Expertenorganisationen besondere Anforderungen an das IT-Management stellen, soll die Beantwortung der Forschungsfragen das theoretische Verständnis über die Wirkungsweise von IT-Governance und IT-Entrepreneurship erweitern. Auf der anderen Seite sollen die Ergebnisse praktische Hinweise liefern, wie das IT-Management im Hinblick auf die Herausforderung der digitalen Transformation in den Krankenhäusern professionalisiert werden kann.

2 Konzeptionelle Entwicklung des Untersuchungsmodells

2.1 Professionalisierungsgrad des IT-Managements

Das IT-Management umfasst alle planerischen, steuernden und überwachenden Aktivitäten, welche im Bezug auf die IT-Ressourcen eines Unternehmens auf strategischer, taktischer und operativer Ebene durchgeführt werden [10]. Nach Weill [16] handelt es sich bei diesen Aktivitäten um die Durchführung IT-bezogener Entscheidungen. Der *Professionalisierungsgrad* des IT-Managements kann über den Umfang und den Formalisierungsgrad der durchgeführten IT-Managementhandlungen definiert werden [10,17]. Professionalität im IT-Management hat demnach eine *quantitative* und eine *qualitative* Dimension.

Zur Definition der quantitativen Dimension können normative Ansätze herangezogen werden [10,18]. Eine zentrale Aufgabe des strategischen IT-Managements ist die Entwicklung einer, mit der Krankenhausstrategie korrespondierenden IT-Strategie [10]. Idealerweise wird die IT-Strategie in eine Finanz- und Investitionsplanung und in ein längerfristiges Projektportfolio überführt [19]. Um den Beitrag der IT zur Erreichung der Unternehmensziele darstellen zu können, erfolgt weiterhin eine strategische Überwachung, bspw. in Form gezielter Evaluationen [10]. Auf taktischer Ebene umfasst das IT-Management alle Aktivitäten, die zur Überführung der IT-Strategie in erfolgreiche IT-Projekte notwendig sind [10]. Hierbei ergeben sich insbesondere Aufgaben rund um die Analyse, Auswahl, Spezifikation, Einführung und Evaluation neuer Systeme [18]. Die Aktivitäten des operativen IT-Managements beziehen sich schließlich auf den erfolgreichen Betrieb aller IT-Komponenten, welche an den klinischen und administrativen Krankenhausabläufen beteiligt sind (Infrastruktur und Netzwerke, Applikationsbetreuung und -wartung, Help- bzw. Servicedesk, etc.) [10].

Die qualitative Dimension des Professionalisierungsgrades kann darüber definiert werden, inwiefern die einzelnen Handlungsweisen formalisiert bzw. standardisiert nach festgelegten Vorgehensweisen durchgeführt werden [12]. Durch die Formalisierung können die IT-Managementaktivitäten einerseits fortlaufend optimiert werden, indem sie an Best-Practice Lösungen ausgerichtet werden [9]. Andererseits ermöglicht die formalisierte Durchführung eine Orientierung der IT-Managementhandlungen an der Unternehmensstrategie [12]. Für die Formalisierung der Managementhandlungen werden in Krankenhäusern sowohl industrielle IT-Governance-Rahmenwerke (bspw. COBIT® oder ITIL®), als auch eigenentwickelte, bzw. generisch entstandene Rahmenwerke genutzt [9].

2.2 IT-Governance

IT-Governance schafft die Rahmenbedingungen für einen zielorientierten Betrieb der unternehmenseigenen IT-Ressourcen [16,20]. Gegenüber dem IT-Management befasst sich IT-Governance nicht mit der Durchführung IT-bezogener Entscheidungen, sondern mit den hierfür notwendigen Befugnissen und Verantwortlichkeiten [16]. IT-Governancemechanismen lassen sich in Strukturen, Prozesse und relationale Ansätze unterteilen [22]. Vor dem Hintergrund der krankenhausspezifischen Vor- und Rahmenbedingungen (vgl. Kap. 1) und im Hinblick auf den Professionalisierungsgrad des IT-Managements erscheinen insbesondere die *hierarchische Positionierung*, das *strategische IT-Alignment* und eine *ausgewiesene IT-Budgetverantwortung* des CIOs bzw. der IT-Abteilung relevante IT-Governancemechanismen darzustellen.

In Krankenhäusern werden abteilungs- und einrichtungsübergreifende Digitalisierungsmaßnahmen vorrangig auf oberen Hierarchiestufen abgestimmt, da sie über die bereichsbezogenen Entscheidungsräume des ärztlichen, pflegerischen und administrativen Bereichs hinausgehen [5,7]. Entsprechend kann eine hohe hierarchische Positionierung des CIOs (bspw. als Mitglied der KHL) die notwendigen Handlungsfreiräume, insbesondere für krankenhausweite Digitalisierungsprojekte, schaffen [9,23,24].

Neben der hierarchischen Einordnung wird die ganzheitlich Plan- und Steuerbarkeit der IT auch von der Ausprägung des strategischen IT-Alignments determiniert [9,10]. Unter IT-Alignment kann die wechselseitige Abstimmung von Strategien, Architekturen, Prozessen sowie Leistungen zwischen IT und (klinischen) Fachabteilungen verstanden werden [19,26]. In Krankenhäusern hat das IT-Alignment in den vergangenen Jahren verstärkt an Bedeutung gewonnen, da sich das Wirkungsfeld des IT-Managements mit zunehmenden informationstechnologischen Potenzialen von dem rein administrativen Funktionsbereich auf den medizinischen und pflegerischen Bereich ausweitet [1,4,10]. Die Synchronisierung von IT und Unternehmensstrategie kann sich zum einen in der Intensität des strategischen Austausches zwischen CIO und KHL und zum anderen in einer, mit der Krankenhausstrategie korrespondierenden IT-Strategie manifestieren [19,26]. Schließlich kann vermutet werden, dass eine ausgewiesene IT-Budgetverantwortung auf Seiten des CIOs bzw. der IT-Abteilung den Professionalisierungsgrad des IT-Managements erhöht, da hierdurch vermutlich langwierige investitionsbezogene Abstimmungsprozesse zwischen den Fachabteilungen verringert werden können [8].

Ausgehend von diesen Vorüberlegungen wird folgende Hypothese aufgestellt:

H1. Je ausgeprägter die IT-Governance, desto höher ist der Professionalisierungsgrad des IT-Managements in den Krankenhäusern.

2.3 IT-Entrepreneurship

Entrepreneurship beschreibt die Neugründung von Organisationen als Reaktion auf identifizierte Marktpotenziale und als Ausdruck spezifischer Gründerpersönlickeiten [27]. Jüngere Definitionen verstehen unter Entrepreneurship das unternehmerische Denken und Handeln innerhalb bestehender Organisation und beziehen sich dabei sowohl auf die Ausprägungen der Organisationskultur [28], als auch auf persönliche Eigenschaften der Organisationsmitglieder [29]. In der vorliegenden Studie wird zwischen *IT-Entrepreneurship-Kultur* und *IT-Entrepreneurship-Persönlichkeit* unterschieden.

Organisationen mit einer hohen IT-Entrepreneurship-Kultur erkennen und erschließen innovative, IT-bezogene Marktpotenziale im Vergleich zu ihren Mitbewerbern vergleichsweise frühzeitig als strategische Wettbewerbsvorteile [28]. Kennzeichnende Merkmale einer IT-Entrepreneurship-Kultur sind eine visionäre und innovative Grundausrichtung, eine vergleichsweise hohe organisatorische Flexibilität und eine Unternehmensleitung, die aktiv den Einsatz innovativer Technologien fördert [28,30]. Eine ausgeprägte IT-Entrepreneurship-Kultur kann sich in Krankenhäusern auf verschiedene Weise positiv auf den Professionalisierungsgrad des IT-Managements auswirken. Zum einen ist die Durchführung strategischer Managementaktivitäten in hohem Maße von der Unterstützung durch die KHL abhängig [8,14]. Andererseits profitiert sowohl die Implementierung als auch der operative Betrieb neuer IT-Lösungen von einer innovativen Grundhaltung und flexiblen Organisationsstrukturen, da hierdurch die Neugestaltung klinischer Arbeitsabläufe nicht nur ermöglicht, sondern im besten Fall auch aktiv von den klinischen Endanwendern eingefordert wird [8,15].

IT-Entrepreneurship-Persönlichkeiten zeichnen sich durch unternehmerisches Denken und Handeln hinsichtlich innovativer IT-Lösungen aus, womit sie nicht selten auch mangelnde Entscheidungsbefugnisse auf mittleren und unteren Hierarchiestufen ausgleichen [27,29]. Eigenschaften von IT-Entrepreneurship-Persönlichkeiten sind eine vergleichsweise hohe Risikobereitschaft, eine eigenständige und proaktive Vorgehensweise sowie eine hohe Partizipationsbereitschaft gegenüber IT-Anwendern [29,31]. In Krankenhäusern kann eine ausgeprägte IT-Entrepreneurship-Persönlichkeit des CIOs den Professionalisierungsgrad des IT-Managements auf verschiedene Weise positiv beeinflussen. Auf der einen Seite können risikoaffine und proaktiv agierende Persönlichkeiten vorherrschende Entscheidungsrestriktionen (vgl. Kap. 1) überwinden und somit zu einem innovativem IT-Betrieb beitragen [29-31]. Auf der anderen Seite kann das IT-Management vermutlich insbesondere in Expertenorganisationen wie Krankenhäusern von einem partizipativen bzw. anwenderorientierten Führungsstil des CIOs profitierten [8,15]. Dies zeigt sich exemplarisch an der Herausforderung, die oftmals hochkomplexen, wissensintensiven medizinischen Behandlungsprozesse in geeigneter Form durch IT zu unterstützen [6]. Gerade klinische Entscheidungen gelten aufgrund des Zusammenspiels von medizinischem Erfahrungswissen, individuellen Patientendaten und medizinischer Evidenzlage als schwer standardisierbar [11]. Entsprechend bedarf es für die Identifikation innovative IT-Lösungen, welche den Klinikern die aktuelle und relevante medizinische Evidenz störungsfrei im Kontext des jeweiligen Behandlungsgeschehens zur Verfügung stellen sollen, nicht nur eine proaktive und risikoaffine Persönlichkeit, sondern auch eine enge Zusammenarbeit zwischen IT-Management und klinischen Experten [6,10,14,18].

Ausgehend von diesen Vorüberlegungen wird folgende Hypothese aufgestellt:

H2. Je ausgeprägter der IT-Entrepreneurship, desto höher ist der Professionalisierungsgrad des IT-Managements in den Krankenhäusern.

2.4 Untersuchungsmodell

Zur Strukturierung des Forschungsvorhabens wurden die konzeptionellen Vorüberlegungen in ein hypothesengeleitetes Untersuchungsmodell überführt (Abb. 1). Die drei zentralen Beschreibungsgrößen des Modells wurden zweistufig angeordnet, da im Hinblick auf den Professionalisierungsgrad des IT-Managements von einem unabhängigen bzw. einem, sich reziprok ergänzendem Einfluss der IT-Governance und des IT-Entrepreneurships ausgegangen wurde. Neben den vorab konzeptionalisierten Beschreibungsgrößen umfasst das Modell Kontrollvariablen als konfundierende Merkmale.

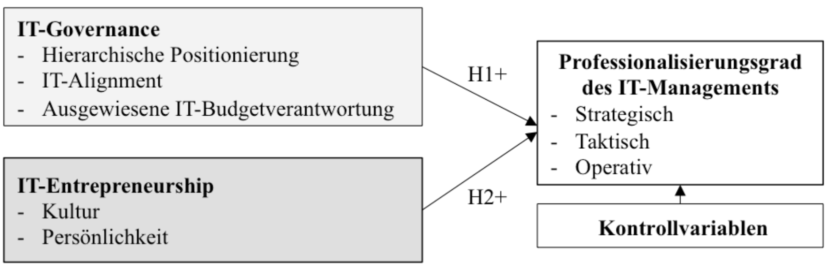


Abbildung 1: Untersuchungsmodell

3 Methode

3.1 Operationalisierung und Datenerhebung

Zur Beantwortung der Forschungsfragen wurde in Anlehnung an Köbler et al. [32] eine empirische Querschnittsuntersuchung als Datenerhebungsmethode ausgewählt. Der hierbei genutzte Fragebogen wurde in mehreren Schritten entwickelt.

In einem ersten, konzeptionellen Schritt wurden die Items für den Fragebogen aus der Literatur abgeleitet [9,10,12,16,18,22,27–31,33–35]. In einem zweiten Schritt wurden die identifizierten Items gemeinsam mit zwei CIOs auf die Übertragbarkeit auf deutsche Krankenhäuser überprüft und teilweise angepasst. Schließlich wurde der Fragebogen in einem Online-Erhebungstool umgesetzt und zwei Pretests unterzogen, in welchen die inhaltliche und technische Plausibilität überprüft wurde. An den Pretests beteiligte sich eine heterogene Gruppe aus vier zusätzlichen CIOs und sechs Wissenschaftlern. Insgesamt wurden 35 Items für die Überprüfung des Untersuchungsmodells entwickelt (vgl. Tab. 2).

Professionalisierungsgrad des IT-Managements: Die quantitative Dimension des Professionalisierungsgrades wurde in Anlehnung an Winter et al. [10] über die Anzahl der durchgeführten IT-Managementaktivitäten operationalisiert. Die qualitative Dimension wurde in Anlehnung an Schlegel et al. [12] über den Formalisierungsgrad der IT-Managementaktivitäten erfasst.

IT-Governance: In Orientierung an vergleichbare Studien [33,36] wurde die Ausprägung der IT-Governance sowohl über organisatorische, als auch über individuelle Eigenschaften operationalisiert. Um die hierarchische Einordnung des IT-Betriebes zu erfassen, wurde gefragt, ob der CIO Mitglied der KHL ist und ob die IT-Abteilung als Stabsstelle fungiert. Das strategische IT-Alignment wurde operationalisiert, indem zum einen die Intensität der strategischen Kommunikation zwischen CIO und KHL und zum anderen das Vorhandensein einer, mit der Krankenhausstrategie korrespondierenden IT-Strategie abgefragt wurden. Schließlich wurde erhoben, ob der CIO bzw. die IT-Abteilung über ein ausgewiesenes IT-Budget verfügen.

IT-Entrepreneurship: Zur Operationalisierung der IT-Entrepreneurship-Kultur wurden in Anlehnung an Bradley et al. [30] drei Items entwickelt, mit denen der Unterstützungsgrad durch die KHL, die Flexibilität der Organisationstruktur und die visionäre Ausrichtung innerhalb des Krankenhauses eingeschätzt werden konnten. Die IT-Entrepreneurship-Persönlichkeit wurde ebenfalls über drei Items operationalisiert, mit denen die proaktive Handlungsweise, die Risikoaffinität und die Partizipationsbereitschaft des CIOs erfasst wurden [31].

Kontrollvariablen: Insgesamt wurden sechs Kontrollvariablen als konfundierende Größen erfasst. Dies waren auf organisatorischer Ebene der Status eines Verbundkrankenhauses, der Status eines Universitätskrankenhauses, die Größe (Bettenzahl) sowie die Trägerschaft des Krankenhauses. Diese Merkmale wurden als positive Einflussgrößen auf die IT-Adoption von Krankenhäusern identifiziert [1,4,37] weshalb vermutet wurde, dass sie auch mit dem Professionalisierungsgrad des IT-Managements zusammenhängen können. Auf Ebene des CIOs wurde in Anlehnung an Smith et al. [36] die Betriebszugehörigkeit des CIOs (in Jahren) sowie in Anlehnung an Burke et al. [38] der akademische Grad (Hochschulabschluss) erhoben. So wurde davon ausgegangen, dass die Erfahrung und Beständigkeit einer Position, sowie die Qualifikation den Professionalisierungsgrad beeinflussen können.

Der Link zu dem Online-Fragebogen wurde in der Feldphase an 1284 CIOs deutscher Krankenhäuser verschickt. Die E-Mailadressen wurden in einer vorgeschalteten Internet- und Telefonrecherche erfasst. Die recherchierten CIOs waren insgesamt für 1675 Krankenhäuser zuständig. Von 305 Krankenhäusern konnten keine Adresse ermittelt werden, da entweder die Position des CIOs nicht existierte, die IT Organisation extern durchgeführt wurde oder in der Recherche keine Auskunft gegeben wurde. Der Erhebungszeitraum erstreckte sich von Februar bis April 2016. In diesem Zeitraum wurden drei Nachfassaktionen durchgeführt.

3.2 Stichprobe

Insgesamt nahmen 188 CIOs an der Umfrage teil, von denen nach einer Vollständigkeits- und Plausibilitätsprüfung 24 Bögen aussortiert wurden. Die verbliebenen 164 auswertbaren Datensätze entsprachen einer Rücklaufquote von 12,8%. In der finalen Stichprobe waren 45,1% der Teilnehmer für mehr als ein Krankenhaus zuständig, sodass die teilnehmenden CIOs insgesamt 397 Krankenhäuser betreuten ($\overline{x}$=2,4; SD=4,6). Die durchschnittliche Betriebszugehörigkeit der CIOs lag bei 11,7 Jahren (SD=7,9). Über einen Hochschulabschluss verfügten 58,5%. Zur Überprüfung der Repräsentativität wurde die Stichprobe in Anlehnung an Köbler et al. [32] nach Größe und Trägerschaft segmentiert und mit der Population deutscher Krankenhäuser verglichen [2] (vgl. Tab.1). Gegenüber der Grundgesamtheit waren kleinere Krankenhäuser mit weniger als 200 Betten leicht unterrepräsentiert und mittlere und größere Krankenhäuser entsprechend überrepräsentiert. Hinsichtlich der Trägerschaft waren private Einrichtungen unterrepräsentiert und öffentliche- sowie freigemeinnützige Krankenhäuser überrepräsentiert.

Tabelle 1: Gegenüberstellung von Stichprobe und Population [2]

Trägerschaft	*unter 200 Betten*	*200 bis 799 Betten*	*800 Betten und mehr*	*Stichprobe*	*Population*
Öffentlich	21,4%	58,9%	19,6%	34,1%	29,7%
Privat	63,2%	26,3%	10,5%	11,6%	35,1%
Frei / Gemeinnützig	44,9%	55,1%	0,0%	54,3%	35,1%
Studienstichprobe	39,0%	53,0%	7,9%	100,0%	
N	64	87	13	164	
Population	56,2%	39,1%	4,7%	100,0%	
N	1113	774	93		1980

In Tabelle 2 werden die deskriptiven Statistiken der Items dargestellt, welche für die Operationalisierung der Beschreibungsgrößen im Untersuchungsmodell genutzt wurden.

Tabelle 2: Deskriptive Statistiken (WB=Wertebereich; n=164)

	Items	WB	$\overline{x}$	SD
strat. IT-Mgmt	Strategische Überwachung (gezielter Evaluationen inkl. Kennzahlenerhebung)[1]	1-3	1,37	0,67
	Erstellung eines Projektportfolios (für ca. 12 Jahre) [1]	1-3	1,55	0,69
	Entwicklung einer mit der Krankenhausstrategie korresp. IT-Strategie[1]	1-3	1,58	0,73
	Strategische Steuerung in Form der Priorisierung und Initiierung von Projekten[1]	1-3	1,76	0,68
	Längerfristige Finanz und Investitionsplanung[1]	1-3	1,77	0,67
takt. IT-Mgmt	Systemevaluation (Informationsbeschaffung, -aufbereitung und -präsentation) [1]	1-3	1,68	0,54
	Systemspezifikation (Beschreibung des SOLL-Zustands, Pflichtenhefts, etc.) [1]	1-3	1,80	0,43
	IT-Projektmanagement (Projektplanung, -begleitung und -abschluss) [1]	1-3	1,85	0,49
	Systemeinführung (Einführungsstrategie, Adaptierung, Mitarbeiterschulung) [1]	1-3	1,88	0,44
	Systemanalyse und -bewertung (bezogen auf den IST-Zustand) [1]	1-3	1,91	0,42
	Systemauswahl (Marktanalyse, Ausschreibung, Angebotsvergleich) [1]	1-3	1,94	0,41
operat. IT-Mgmt	Durchführung eines IT-bezogenen Rechnungswesen[1]	1-3	1,91	0,69
	Durchführung eines IT-bezogenen Vertragsmanagement[1]	1-3	1,93	0,70
	Schulungen bzw. Trainings klinischer Endanwender[1]	1-3	2,04	0,64
	Betrieb des Helpdesk / Servicedesk[1]	1-3	2,14	0,59
	Applikationsbetreuung und -wartung[1]	1-3	2,21	0,51
	Steuerung und Überwachung von Infrastruktur und Netzwerken[1]	1-3	2,23	0,50
IT-Governance	Intensität strategische Kommunikation[2]	0-10	0,05	0,23
	Existenz einer, mit der Krankenhausstrategie korrespondierenden IT-Strategie[3]	0-1	0,18	0,38
	CIO ist Mitglied der Krankenhausleitung[3]	0-1	0,05	0,23
	IT-Abteilung ist als Stabsstelle eingesetzt[3]	0-1	0,18	0,38
	CIO verfügt über ein ausgewiesenes IT-Budget[3]	0-1	0,26	0,44
	IT-Abteilung verfügt über ein ausgewiesenes IT-Budget[3]	0-1	0,76	0,43
IT-Entrep.	Als CIO muss ich mich intensiv mit den Bedürfnissen der Anwender befassen[4]	1-4	2,54	0,84
	Als CIO arbeite und entscheide ich weitestgehend selbstbestimmt[4]	1-4	2,29	0,66
	Als CIO muss ich mich auch auf neue, nicht bewährte Lösungen einlassen[4]	1-4	1,98	0,59
	Unsere Krankenhausleitung fördert aktiv innovative IT-Lösungen[4]	1-4	2,68	0,80
	Unser Krankenhaus ist bzgl. des Einsatzes innovativer IT sehr flexibel[4]	1-4	2,77	0,78
	In unserem Krankenhaus herrscht eine Zukunftsvision, die auch die IT umfasst[4]	1-4	2,81	0,80

[1] 1=„keine Durchführung“; 2=„nicht-formalisierte Durchführung“; 3=„formalisierte Durchführung in Anlehnung an eigens entwickelte oder industrielle IT-Governancerahmenwerke“,
[2] Anzahl von häufig zw. CIO und KHL ausgetauschter, strategischer Informationen. Mehrfachauswahl von zehn Informationstypen, z.B. „IT-relevante Krankenhausziele“, „Sicherheitskonzept“, „Prozesse“
[3] 0=„Nein“; 1=„Ja“
[4] 1=„Stimme überhaupt nicht zu“; 2=„Stimme eher nicht zu“; 3=„Stimme eher zu“; 4=„Stimme voll und ganz zu“

3.3 Datenanalyse

Um die Annahmen des Untersuchungsmodells zu überprüfen, wurde eine multiple Regressionsanalyse mit SPSS23® durchgeführt. Multiple Regressionsanalysen eignen sich als hypothesenprüfende Verfahren, da sie den Einfluss mehrerer unabhängiger Variablen (Prädiktoren) auf eine abhängige Variable (Kriterium) testen [39].

Als Kriterium diente der *Professionalisierungsgrad des IT-Managements*. Zur Quantifizierung des Professionalisierungsgrades wurde ein gewichteter Summenscore gebildet, indem pro durchgeführte IT-Managementaktivität ein Punkt und pro formalisiert durchgeführter IT-Managementaktivität 1,5 Punkte vergeben wurden.[1] Der gewichtete Summenscore wurde skaliert, sodass der Professionalisierungsgrad des IT-Managements in einem Wertebereich von 0 bis 100 Punkten lag.

Um das Untersuchungsmodell zu überprüfen, wurden fünf Merkmale zur Beschreibung der IT-Governance[2], zwei Merkmale zur Beschreibung des IT-Entrepreneurships sowie sechs Kontrollvariablen als konfundierende Größen in das Modell eingeschlossen (vgl. Kap. 3.1). Zur Quantifizierung der *IT-Entrepreneurships-Kultur* und *-Persönlichkeit* wurden Summenscores gebildet, indem jeweils der Zustimmungsgrad zu den drei aufgestellten Aussagen (vgl. Tab. 2) aufaddiert wurde.

Für die statistischen Analysen wurden fehlende Werte durch Mittelwerte ersetzt. Zur Überprüfung der Modellvoraussetzungen wurde auf Homoskedastizität sowie auf Normalverteilung der Residuen getestet. Zur Prüfung auf Multikollinearität wurden Toleranzwerte sowie der Variance Inflation Factor (VIF) berechnet [39]. Zudem wurden die signifikant in dem Modell verbliebenen Prädiktoren korreliert.

4 Ergebnisse

Von maximal 100 Punkten, die für den Professionalisierungsgrad des IT-Managements erreicht werden konnten, erzielten die befragten Einrichtungen durchschnittlich 42 (SD=14; n=164). Der niedrigste Wert lag bei 15, der höchste bei 100 Punkten. Nur knapp jedes fünfte teilnehmende Krankenhaus erreichte einen Wert über 50 (19,5%; n=164). Strategische IT-Managementaktivitäten wurden gegenüber operativen und taktischen Managementaktivitäten vergleichsweise selten durchgeführt. Dies galt insbesondere für die strategische Überwachung, aber bspw. auch für die Entwicklung von längerfristigen IT-Projektportfolios und für die IT-Strategieplanung (vgl. Tab. 2). Sowohl im strategischen, als auch im taktischen und operativen Bereich wurden evaluierende bzw. überwachende IT-Managementhandlungen im Gegensatz zu planenden und ausführenden Aktivtäten seltener durchgeführt (vgl. Tab. 2).

In Tabelle 3 werden die Koeffizienten der Prädiktoren im Regressionsmodell, gegliedert in Anlehnung an die Operationalisierung der Beschreibungsgrößen (vgl. Kap. 3.1), dargestellt. Von den 13 Prädiktoren, welche in das Modell eingeschlossen wurden, ergaben sich neun signifikante Beta-Koeffizienten (vgl. Tab. 3). Zusammengenommen erklärten die eingeschlossenen Prädiktoren 47,0% der Varianz des Professionalisierungsgrades (korr. R^2). Drei der fünf überprüften IT-Governancemerkmale zeigten einen signifikant positiven Einfluss auf den Professionalisierungsgrad. Ebenfalls wirkte sich die IT-Entrepreneurship-Kultur und -Persönlichkeit signifikant posi-

[1] Die moderierende Gewichtung folgt der Annahme, dass formalisiert durchgeführte Managementaktivitäten im Hinblick auf den Professionalisierungsgrad nicht zwangsläufig doppelt so hoch gewertet werden können, wie ad-hoc durchgeführten Aktivitäten, da die Aktivität an sich der wesentliche Aspekt ist.

[2] Die Variable “CIO ist Mitglied der KHL“ wurde aufgrund der ungleichen Verteilung der Merkmalsausprägungen (5% zu 95%) aus der Regressionsanalyse herausgenommen.

tiv auf den Professionalisierungsgrad aus. Auch der Status eines Universitätsklinikums und der des Verbundkrankenhauses zeigten einen signifikant positiven Einfluss.

Die Residuen waren normalverteilt und Homoskedastizität lag nicht vor, sodass die Modellvoraussetzungen erfüllt waren. Die errechneten Korrelationsstatistiken gaben keine Hinweise auf Multikollinearität (vgl. Tab. 3). In Tabelle 4 werden die Korrelationen der signifikanten Prädiktoren in Form einer Korrelationsmatrix dargestellt. Die Korrelation der Prädiktoren weisen auf schwache, positive Korrelationen zwischen einzelnen Prädiktoren hin, wobei die „Intensität der strategischen Kommunikation" mit drei signifikanten Korrelationskoeffizienten am häufigsten korrelierte (vgl. Tab. 4).

Tabelle 3. Koeffizienten der Prädiktoren im Regressionsmodell (n=164)

Prädiktor	*Beta*	*Sig.*	*Toleranz*	*VIF*
Intensität der strat. Kommunikation	0,205	0,002	0,798	1,254
Existenz eines IT-Strategie	0,062	0,329	0,803	1,246
IT-Abteilung ist Stabsstelle	0,030	0,625	0,877	1,140
CIO verfügt über IT-Budget	0,231	0,000	0,933	1,071
IT-Abteilung verfügt über IT-Budget	0,173	0,004	0,906	1,104
IT-Entrepreneurship-Kultur	0,192	0,004	0,841	1,150
IT-Entrepreneurship-Persönlichkeit	0,224	0,000	0,893	1,120
Status eines Universitätskrankenhauses	0,147	0,210	0,822	1,217
Status eines Verbundkrankenhauses	0,247	0,000	0,902	1,108
Trägerschaft (Privat)	-0,850	0,153	0,939	1,064
Größe (Bettenzahl)	0,760	0,271	0,682	1,467
Betriebszugehörigkeit (in Jahren)	-0,187	0,002	0,916	1,092
CIO hat Hochschulabschluss	0,100	0,094	0,924	1,082

Tabelle 4. Korrelation der signifikanten Prädiktoren (*p<0,05; n=164)

		1	2	3	4	5	6	7	8
1	Intensität strat. Kommunikation	1,000	.	.	.	.	.	.	.
2	CIO verfügt über IT-Budget	0,131	1,000	.	.	.	.	.	.
3	IT-Abt. verfügt über IT-Budget	0,149	0,113	1,000	.	.	.	.	.
4	IT-Entrep.-Kultur	0,377*	0,048	0,254*	1,000	.	.	.	.
5	IT-Entrep.-Persönlichkeit	0,205*	0,101	0,127	0,172*	1,000	.	.	.
6	Universitäts-krankenhaus	0,177*	0,196*	0,085	0,058	0,199*	1,000	.	.
7	Verbund-krankenhaus	-0,052	0,028	-0,014	-0,189*	0,036	-0,198*	1,000	.
8	Betriebs-zugehörigkeit	-0,057	0,024	-0,012	0,153	0,096	-0,147	0,079	1,000
	Hochschul-Abschluss	0,082	0,108	0,098	-0,065	0,037	0,214*	0,029	-0,098

5 Diskussion

Ein professionell agierendes IT-Management gilt als Dreh- und Angelpunkt für eine erfolgreiche Umsetzung der digitalen Transformation [4,10,19,32]. Vor diesem Hintergrund erscheint es umso bemerkenswerter, dass bis dato keine empirisch fundierten Erkenntnisse darüber existieren, wie professionell das IT-Management in Krankenhäusern tatsächlich ist und wodurch der Professionalisierungsgrad determiniert wird. Um diese Forschungslücke zu schließen, wurde in der vorliegenden Studie ein hypothesengeleitetes Untersuchungsmodell entwickelt und anhand der Daten von 164 CIOs überprüft.

Die Ergebnisse der Studie deuten auf diverse Professionalisierungspotenziale des IT-Managements hin. So müsste ein Großteil der befragten Krankenhäuser die IT-Strategieplanung intensivieren und entsprechende IT-Vorhaben konsequent in adäquate Finanz- und Investitionsplanungen sowie in längerfristige Projektportfolios überführen, wenn Digitalisierungspotenziale proaktiv genutzt werden sollen [10,19]. Zudem müsste der Einsatz von Evaluierungsmaßnahmen wie bspw. Systemanalysen, Anwenderbefragungen oder IT-Benchmarks verstärkt fokussiert werden, damit der Digitalisierungsfortschritt fortlaufend überwacht und sein Wertbeitrag transparent dargestellt werden kann [9,10,13].

Bei entsprechenden Professionalisierungsbemühungen sieht sich das IT-Management der Krankenhäuser mit unterschiedlichen Herausforderungen konfrontiert [5,7-9,12-14]. Vor diesem Hintergrund wurde die Annahme getroffen, dass eine ausgeprägte IT-Governance sowie ein hoher IT-Entrepreneurship den Professionalisierungsgrad positiv beeinflussen. Die Ergebnisse der Studie konnten diese Annahmen weitestgehend bestätigen. So scheint sich der intensive Austausch zwischen CIO und KHL positiv auf den Professionalisierungsgrad des IT-Managements auszuwirken. Dies gilt auch für eine ausgewiesene IT-Budgetverantwortung des CIOs bzw. der IT-Abteilung. Der Einfluss der hierarchischen Positionierung des CIOs konnte aus methodischen Gründen nicht überprüft werden. Jedoch zeigten die Ergebnisse, dass das IT-Management in den meisten Krankenhäusern auf mittleren und unteren Hierarchiestufen eingeordnet ist. Die Wirkungsweise der betrachteten IT-Governancemechanismen sollten im Zusammenhang mit dem ebenfalls betrachteten IT-Entrepreneurship interpretiert werden. So deuten die Ergebnisse darauf hin, dass der Professionalisierungsgrad des IT-Managements durch eine ausgeprägte Unterstützung der KHL positiv beeinflusst wird. Darüber hinaus scheint ein professionell agierendes IT-Management mit einer visionären Grundhaltung und flexiblen Organisationsstrukturen einherzugehen. Schließlich zeigten die Ergebnisse, dass sich eine ausgeprägte Entrepreneurship-Persönlichkeit des CIOs positiv auf den Professionalisierungsgrad des IT-Managements auswirkt. So kann vermutet werden, dass CIOs auf mittleren und unteren Hierarchiestufen mangelnde Entscheidungsbefugnisse durch unternehmerisches Denken und Handeln ausgleichen [27,29]. Darüber hinaus scheint das IT-Management insbesondere in wissensintensiven Expertenorganisationen wie Krankenhäusern von einem anwenderorientierten Führungsstil des CIOs zu profitieren [8,15].

Die vorliegende Studie liefert erste Hinweise auf wesentliche Begleitumstände und Vorbedingungen eines professionellen IT-Managements. Hierdurch erhalten Krankenhäuser empirisch fundierte Hinweise für eine erfolgreiche Umsetzung der digitalen Transformation. Insbesondere die Rolle einer ausgeprägten Entrepreneurship-Persönlichkeit des CIOs kann einen Ansatz für weitergehende Forschungsarbeiten liefern. In einem ersten Schritt könnte eine reliable und valide Operationalisierung des Konstrukts, bspw. durch faktoranalytische Verfahren fokussiert werden. Weiterhin könnten Folgestudien die Wechselwirkung von IT-Governance und IT-Entrepreneurship durch Interaktionstests näher spezifizieren. Auch wäre es von Interesse, die IT-Performance der Krankenhäuser in einem erweiterten, mehrstufigen Untersuchungsmodell als zusätzliche Zielgröße zu berücksichtigen. Die IT-Performance könnte dabei über die IT-Unterstützung klinischer Prozesse operationalisiert werden. Schließlich könnte die Übertragbarkeit der Ergebnisse auf andere Betätigungsfelder des IT-Managements, insbesondere auf das Feld wissensintensiver Expertenorganisationen (z.B. Hochschulen) überprüft werden.

6 Limitation

Bei der Interpretation der Studienergebnisse müssen mehrere Limitationen berücksichtigt werden. Zum einen steht die betrachtete Stichprobe nicht repräsentativ für die Grundgesamtheit der deutschen Krankenhäuser. Kleinere und nicht private Einrichtungen waren unterrepräsentiert. Die Faktoren zur Beschreibung des Professionalisierungsgrades und des Entrepreneurships wurden zwar inhaltlich validiert, jedoch nicht auf Reliabilität überprüft. Die vorgenommene, moderierende Gewichtung folgte zudem ausschließlich inhaltlichen Erwägungen. Weiterhin wurde die IT-Governance größtenteils über binäre Merkmale operationalisiert. Zukünftige Ansätze sollten hier, komplementär zur Operationalisierung von IT-Entrepreneurship, validierte Item-Skalen nutzen. Zwischen den einzelnen Prädiktoren zeigten sich schwache, jedoch signifikante positive Korrelationen. Obwohl dies die Interpretierbarkeit der Ergebnisse im Hinblick auf die zusätzlich berechneten Multikollinearitätstatistiken nicht schwächt, sollten zukünftige Ansätze eine überschneidungsfreie Operationalisierung adressieren. Auch aus der Nutzung des gewählten Analyseverfahrens ergeben sich Limitationen. Regressionsanalysen eignen sich insbesondere für die Überprüfung prozesshafter Modelle mit einseitigen Abhängigkeiten zwischen den Merkmalen. Diese kausalen Beziehungen spiegeln die Realität des betrachteten Untersuchungsfeldes jedoch nur bedingt wider, da sich IT-Governance und das IT-Management vermutlich rekursive beeinflussen. Schließlich berücksichtigt das Modell keine Strukturbrüche zwischen den individuellen Merkmalen des CIOs und den Merkmalen der Organisation. Zukünftige Ansätze sollten daher auf komplexere Analyseverfahren, wie bspw. Mehrebenenmodelle oder Strukturgleichungsverfahren zurückgreifen.

Literaturquellen

1. Agarwal, R., Gao, G., DesRoches, C., Jha, A.K.: Research Commentary —The Digital Transformation of Healthcare. Current Status and the Road Ahead. Inform. Syst. Res. 21, 796–809 (2010)

2. Gesundheitsberichterstattung des Bundes, http://www.gbe-bund.de/ (Abgerufen am: 26.08.2016)
3. Haas, P.: Gesundheitstelematik. Grundlagen, Anwendungen, Potenziale. Springer Berlin Heidelberg (2006)
4. Buntin, M.B., Burke, M.F., Hoaglin, M.C., Blumenthal, D.: The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff. 30, 464–471 (2011)
5. Genzel, H., Siess, M.: Ärztliche Leistungs- und Organisationsstruktur im modernen Krankenhaus - Zu den zukünftigen Herausforderungen für das ärztliche Krankenhausmanagement aus medizinischer, rechtlicher und ökonomischer Sicht. Medizinrecht (1999)
6. Jobst, F.: IT zur Prozessgestaltung im Krankenhaus – Wie bekommt man die optimale Kombination von IT-Anwendungen? In: Schlegel, H. (ed.) Steuerung der IT im Klinikmanagement, pp. 225–251. Vieweg+Teubner, Wiesbaden (2010)
7. Leuzinger, A., Luterbacher, T.: Mitarbeiterführung im Krankenhaus. Spital, Klinik und Heim. Huber, Bern u.a. (2000)
8. Avgar, A.C., Litwin, A.S., Pronovost, P.J.: Drivers and barriers in health IT adoption: a proposed framework. Appl. Clin. Inform. 3, 488–500 (2012)
9. Thatcher, M.: IT Governance in Acute Healthcare: A Critical Review of Current Literature. In: George, C., Whitehouse, D., Duquenoy, P. (eds.) eHealth: Legal, Ethical and Governance Challenges, pp. 349–370. Springer, Berlin, Heidelberg (2013)
10. Winter, A., Haux, R., Ammenwerth, E., Brigl, B., Hellrung, N., Jahn, F.: Health Information Systems. Springer London, London (2011)
11. Lenz, R., Reichert, M.: IT support for healthcare processes – premises, challenges, perspectives. Data & Knowledge Engineering 61, 39–58 (2007)
12. Schlegel, H.: IT-Governance mit COBIT® - Methodenunterstützung für das Management. In: Schlegel, H. (ed.) Steuerung der IT im Klinikmanagement, 1, pp. 7–27. Vieweg+Teubner, Wiesbaden (2010)
13. Kutscha, A., Kutscha, U.: Die Balanced Scorecard als Management- und Controllinginstrument – Nutzenpotentiale für die IT im Krankenhaus. In: Schlegel, H. (ed.) Steuerung der IT im Klinikmanagement, pp. 53–71. Vieweg+Teubner, Wiesbaden (2010)
14. Spetz, J., Keane, D.: Information technology implementation in a rural hospital: a cautionary tale. J. Healthc. Manag. 54, 337-47; discussion 348 (2009)
15. Cresswell, K., Sheikh, A.: Organizational issues in the implementation and adoption of health information technology innovations: an interpretative review. Int. J. Med. Inform. 82, e73-86 (2013)
16. Weill, P.: Don't Just Lead Govern: How Top-Performing Firms Govern IT. MIS Quart. 3, 1–17 (2004)
17. Schein, E.H.: Organizational Socialization and the Profession of Management. Industrial Management Review (1968)
18. Ammenwerth, E., Haux, R., Knaup-Gregori, P., Winter, A.: IT-Projektmanagement im Gesundheitswesen. Lehrbuch und Projektleitfaden Taktisches Management von Informationssystemen. Schattauer, Stuttgart (2015)
19. Bush, M., Lederer, A.L., Li, X., Palmisano, J., Rao, S.: The alignment of information systems with organizational objectives and strategies in health care. Int. J. Med. Inform. 78, 446–456 (2009)
20. De Haes, S., Van Grembergen, W.: IT governance and its mechanisms. Information Systems Control Journal 1, 27–33 (2004)
21. Peterson, R.: Crafting Information Technology Governance. Inform. Syst. Manage. 21, 7–22 (2004)

22. Weill, P., Ross, J.: A matrixed approach to designing IT governance. Sloan. Manage. Rev. 46, 25–35 (2005)
23. Raghunathan, B., Raghunathan, T.: Relationship of the rank of informationsystems executive to the organizational role and plan-ningdimensions of information systems. J. Manage. Inf. Syst. 6, 111–126 (1989)
24. Applegate, L.M., Elam, J.J.: New Information Systems Leaders. A Changing Role in a Changing World. MIS Quart. 16, 469–489 (1992)
25. Nahapiet, J., Ghoshal, S.: Social Capital, Intellectual Capital, and the Organizational Advantage. Acad. Manage. Rev. 23, 242–266 (1998)
26. Reich, B.H., Benbasat, I.: Factors That Influence the Social Dimension of Alignment between Business and Information Technology Objectives. MIS Quart. 24, 81 (2000)
27. Freiling, J., Gersch, M.: Auf dem Weg zu einer „Dienstleistungstheorie“: das Zusammenspiel individueller und kollektiver Fähigkeiten im Kontext aktueller Theorieentwicklungen. In: Freiling, J. (ed.) Wirkungsbeziehungen zwischen individuellen Fähigkeiten und kollektiver Kompetenz, pp. 99–130. Hampp, München (2008)
28. Russell, R.D.: How Organisational Culture Can Help to Institutionalise the Spirit of Innovation in Entrepreneurial Ventures. Journal of OrgChange Mgmt. 2, 7–15 (1989)
29. Heinze, K.L., Weber, K.: Toward Organizational Pluralism. Institutional Intrapreneurship in Integrative Medicine. Organ. Stud. (2015)
30. Bradley, R.V., Byrd, T.A., Pridmore, J.L., Thrasher, E., Pratt, R.M.E., Mbarika, V.W.A.: An empirical examination of antecedents and consequences of IT governance in US hospitals. J. Inf. Technol. 27, 156–177 (2012)
31. Patterson, F., Máire, K. and Geraldine, G.-R.: Characteristics and behaviours of innovative people in organisations. Literature Review prepared for the NESTA Policy & Research Unit, 1–63 (2009)
32. Köbler, F., Fähling, J., Krcmar, H., Jan Marco, L.: IT-Governance und IT-Entscheidertypen in deutschen Krankenhäusern - Eine empirische Untersuchung unter Krankenhaus-IT-Leitern. Wirtschaftsinformatik 52, 353–365 (2010)
33. Leidner, D.E., Preston, D., Chen, D.: An examination of the antecedents and consequences of organizational IT innovation in hospitals. J. Strateg. Inf. Syst. 19, 154–170 (2010)
34. Heinrich, L.J., Riedl, R., Stelzer, D.: Informationsmanagement. Grundlagen, Aufgaben, Methoden. Oldenburg Wirtschaftsverlag, München (2014)
35. Croteau, A.M., Bergeron, F., Dubsky, J.: Contractual and Consensual Profiles for an Interorganizational Governance of Information Technology. International Business Research 6 (2013)
36. Smith, A.L., Bradley, R.V., Bichescu, B.C., Tremblay, M.C.: IT Governance Characteristics, Electronic Medical Records Sophistication, and Financial Performance in U.S. Hospitals. An Empirical Investigation. Decis. Sci. 44, 483–516 (2013)
37. Black, A.D., Car, J., Pagliari, C., Anandan, C., Cresswell, K., Bokun, T., McKinstry, B., Procter, R., Majeed, A., Sheikh, A.: The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med. 8, e1000387 (2011)
38. Burke, D., Menachemi, N., Brooks, R.: Health care CIOs: assessing their fit in the organizational hierarchy and their influence on information technology capability. Health Care Manag. 25, 167–172 (2006)
39. Fahrmeir, L., Kneib, T., Lang, S.: Regression. Modelle, Methoden und Anwendungen. Springer, Berlin u.a. (2009)

Introducing Archetype Theory to Information Systems Research: A Literature Review and Call for Future Research

Raphael D. Schilling, M. Kazem Haki, and Stephan Aier

University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland
{raphael.schilling,kazem.haki,stephan.aier}@unisg.ch

Abstract. Studying organizational configurations on the one hand and the dynamics of organizational change on the other hand are dominant topics of interest in the information systems (IS) discipline. Studies in each of these research streams take advantage of various well-established theoretical lenses from reference disciplines such as management science. In this study, we take a closer look at archetype theory, which combines these two research streams and which eventually provides a dynamic perspective on organizational configurations. Through a literature review, this study provides a comprehensive understanding of archetype theory (i.e., its constitutive constructs and assumptions) as well as on its application in studying dynamics of configurations. In introducing archetype theory to IS research, we discuss the explanatory power of the respective theory for investigating IS phenomena as well as the methodological and theoretical implications of employing the theory in IS research.

Keywords: Archetype Theory, Literature Review, Change, Configuration

1 Introduction

Investigating configurations has long been central to management research [1-3] as well as to information systems (IS) research [4, 5]. Respective studies seek for patterns, classifications, and sets of structures that differ in their fundamental characteristics and that are effective under different circumstances [e.g., 3, 6]. Such configurations have been researched for IS governance [e.g., 4, 7, 8], ERP implementation [e.g., 9, 10], inter-organizational IS [e.g., 11], among others. However, as fast-moving change has become the natural mode of organizational life [12], the existing *static and deterministic view* on configurations would not appropriately account for the dynamics of change in a turbulent environment [6]. Consequently, configurations should also be examined from the *perspective of change* as they emerge from the dyadic, dynamic interactions between organizations and their ever-changing environment [6, 13].

In the extant IS literature, scholars dominantly examined *either* configurations [e.g., 4, 5, 7, 11], *or* dynamics of change [e.g., 14, 15-17]. As the use of proper theoretical lenses guides scholars in both theory building and theory testing in IS research [18-23],

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Schilling, R.D.; Haki, M.K.; Aier, S. (2017): Introducing Archetype Theory to Information Systems Research: A Literature Review and Call for Future Research, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 574-588

in this study we promote the use of archetype theory as a purposeful theoretical lens to *simultaneously* study configurations and dynamics, i.e., the configurations that result from dynamics of change. Compared to other frequently used reference theories in IS, such as institutional theory [24, 25], *archetype theory* has not gained recognition in IS research.

After generating their theory on the information technology (IT) function in organizations, Guillemette and Paré [26] in the discussion of their resulted insights state that the lens of archetype theory guides theory-driven investigation of configurations and their inherent dynamics, thus opens new avenues for studying various IS phenomena. As such, with the goal of introducing archetype theory to IS research and commencing preliminary discussions on its implications in IS research, in this study we seek for (i) providing a comprehensive understanding of archetype theory and its underlying analytical constructs and assumptions, and (ii) reviewing how and for which purposes extant research has applied archetype theory. Therefore, compared to existing reviews in management research that focus on archetypes in, for instance, professional service firms [27, 28] and sports organizations [29], the focus of our review is the theory itself and its applications. To this end, our study uncovers the explanatory power of archetype theory in investigating a wide range of IS phenomena, since it caters an in-depth and profound understanding of the underlying mechanisms of change as well as the configurations resulting from change. Finally, we discuss methodological and theoretical implications of employing archetype theory for IS research.

2 Archetype Theory

The identification of optimal configurations and their dynamics have long been the focus of research in management studies [30]. Research on this topic evolved through three schools of thought, namely gradualist, contingency, and structural adjustment paradigms [31]. During the development of the classical management theories in 1950s, academia believed in "one best way" to structure organizations [32]. The *gradualist paradigm* consequently propagates evolutionary approaches [33], similar to Darwin's model of evolution and is reflected in, for instance, lifecycle and maturity model metaphors. Due to neglecting the context in which configurations are embedded, this initial understanding was challenged by other theorists [41], focusing on the contextual factors, which ultimately lead to the *contingency paradigm*. The latter brings a central argument to the forefront of configuration research: "the external circumstances that produce particular organizational designs, and the idea that there is an appropriate linkage between the external, the internal, and performance" [32, p. 400]. Later, institutional theory adds to the contingency paradigm by revising the linkage of internal and external contingencies and performance through a stronger focus on institutional pressures [32]. As such, organizations are understood as reflections or responses to rules, beliefs, and conventions in their surrounding environment [34]. However, organizations have been evolving differently, even when exposed to the very same institutional pressures [35]. Theories of the *structural adjustment paradigm* thus focus on the dynamics of organizational adaptation and explain the movement between different organizational configurations in the same organizational context. In this paradigm, configuration theory has been influential in taking punctuated equilibrium assumptions (i.e., necessity of ongoing fit between contingency factors and configuration parameters) and in considering a significant number of contingencies resulting in rich descriptions of optimal configurations [36]. Configuration theory posits that configurations can be determined through typologies, taxonomies, and archetypes [27]. As the notion of archetype in configuration theory tries to not only consider optimal configurations but also the underpinning mechanisms that bring about these configurations, it has further been recognized as archetype theory[1]. This theory comprises two key aspects: (i) the *exploration of organizational archetypes* as well as (ii) the *analysis of change*. The following sub-sections describe these aspects in detail.

2.1 Exploration of Organizational Archetypes

Greenwood, Hinings, and Laughlin [35, 38, 39] elaborated the concept of optimal configurations through the notion of archetype and archetype theory respectively. According to archetype theory, an archetype comprises the twin concepts of *interpretative scheme* and *structural arrangement*:

The *interpretative scheme* describes an organization's conception on what it should be doing, how it should be doing, and how it should be judged. This conception is shaped by the prevailing set of ideas, beliefs, and values [38]. The *structural arrangement* implements and reinforces the ideas, beliefs, and values through establishing organizational structures and processes that reflect the respective beliefs and values [38]. Thus, there is a strong interrelation between the interpretative scheme and the structural arrangement as they reinforce each other.

In an ideal case, organizations will evolve towards a situation of organizational *coherence*, where the structural arrangement and the interpretative scheme represent an "appropriate design for adequate performance" [38, p. 295]. As such, in the coherence situation, interpretive scheme and structural arrangement are in line with each other and represent a specific archetype. However, neither do all organizations change in the same way nor will all of them reach a level of high performance. Thus, the analysis of change is an integral part of archetype theory, for which the concept of *change tracks* has been applied.

2.2 Analysis of Change

The identification of archetypes is a preparatory step for the explanation of change. By the identification of the archetype, an organization is situated in one of the following *positions* [38]:

1. *Archetype coherence*, where the interpretative scheme and structural arrangement match and thus reflect and reinforce each other.

[1] Theorists introduced archetype theory both as a subordinate of configuration theory [27] and as a synonym of configuration theory [37].

2. *Embryonic archetype coherence,* where some design elements are discordant as interpretative scheme and structural arrangement do not perfectly match.
3. *Schizoid incoherence*, where organizations show the presence of two different archetypes at the same time and thus competing interpretative schemes and structural arrangements.

To explain the actual change process, archetype theory outlines the movement of organizations along the abovementioned positions through the concept of *tracks* [38]:

1. *Track A - Inertia:* Most organizations will stick to one archetype for a lengthy period of time. This track describes a situation of archetype coherence with incremental changes, where only slight structural adjustments within a particular archetype can be observed.
2. *Track B - Aborted excursion:* Here, organizations shift from a position of archetype coherence towards an embryonic archetype coherence and back to archetype coherence.
3. *Track C - Reorientation:* Describes the typical transformation situation, where organizations move from an archetype to another. This includes fundamental changes in both the structural arrangement and interpretative scheme.
4. *Track D - Unresolved excursion:* Describes a failed change process. The organization is trapped between two competing archetypes.

Relying on archetype theory's focus as well as on its well-defined constructs and relations, it has a considerable potential to help scholars understand the dynamics of change and its resultant configurations (i.e., archetypes) in IS research. The concept of organizational coherence considers both, tangible artifacts such as structures and processes but also intangibles such as values and beliefs. It also acknowledges the role of dynamics of change by defining multiple change tracks. Therefore, it is worthwhile analyzing how archetype theory has already been applied and, more specifically, how configurations and change have been examined through this theory in the extant literature.

3 Research Method

Despite its potential, archetype theory has not gained much attention in IS research yet [31]. We therefore opt for a literature review, which is considered suitable to identify potential implications for prospective research [40]. This section describes how we identified and analyzed the relevant literature.

3.1 Literature Selection and Review Process

In order to identify prior research relevant to archetype theory, we searched for articles, containing "archetype theory"[2] in either the title, abstract, or keywords. To extract high-quality contributions, we limited our search process on scientific databases to peer-reviewed, scholarly journals (with no limitation on the publication date and the type of journal) and excluded other types of publications (e.g., books, projects, conference proceedings). We identified the related articles by scanning Business Source Premier as well as ProQuest scientific databases. The goal was to cover a wide range of research disciplines, including IS. Not surprisingly, we did not find any journal article in IS, we thus extended the search process and further included AIS-supported conference articles in IS, which were retrieved from the AIS Electronic Library.

We subsequently went carefully through the abstracts of all articles and excluded those that were not relevant (see **Table 1**).

Table 1. Search and selection process of the literature review

Database	Business Source Premier	ProQuest	AIS Electronic Library
Search Term	"archetype+theory" IN title OR abstract OR keywords		"archetype+theory"
Filter	Only look for scholarly journals		none
# of articles found *(total: 42)*	7	34	1
# of articles excluded and reasons for exclusion *(total: 26)*	• Article has no relation to archetype theory according to abstract (10 articles) • Article is referring to "jungian archetype theory", which is different than the archetype theory in management science (9 articles) • Article is a book review (1 article) • Article is written in other languages than English (1 article) • Article is a seminal article on archetype theory (5 articles)		
# of articles considered for review *(total: 16)*	16 application articles that employed archetype theory as a theoretical lens in investigating their phenomena of interest.		

We also differentiated seminal articles (contributing to the seminal assumptions and constructs of archetype theory) from application articles (applying assumptions and constructs of archetype theory to their phenomenon of interest). Seminal articles are used for the construction of the analysis framework, which is described in the following section. We used the analysis framework to code the application articles to gain insights on how scholars applied archetype theory in their respective research.

[2] It is noteworthy that we did not search for the term "archetype" alone or other relevant terms such as "configuration" and "gestalt" [41]. This is due to the focus and scope of our research in reviewing archetype theory itself and its applications, not in identifying the derived archetypes or configurations in the extant research.

3.2 Analysis Framework

Following the guidelines of Webster and Watson [40] and Fettke [42], we developed an analysis framework to guide literature analysis. This analysis framework comprises the *constructs* of archetype theory (as introduced in Section 2) as well as further components to gain insights on how archetype theory has been *applied*. The constitutive components of the analysis framework are:

Use of the theory: Archetype theory can be used to explore archetypes and/or to analyze change. We classified, whether new archetypes are explored and whether changes are analyzed based on archetype theory. This helps us gain insight into the purposes archetype theory is dominantly used for.

Exploration of the archetype: This component of the analysis framework aims at identifying how archetypes have been described. We therefore extracted all the proposed archetypes along with the respective interpretive scheme and structural arrangement of each archetype.

Change: Explaining the dynamics of change is one of the fundamental premises of archetype theory. Therefore, this component of the analysis framework captures the core findings of articles on change, particularly with regard to different types of change tracks.

Research method: In order to capture the dynamics of archetypes, researchers need to employ appropriate methods to investigate structural arrangements and, even more challenging, interpretive schemes. This will support future research in selecting appropriate research methods. Therefore, in the analysis framework, the employed research methods are classified into conceptual, quantitative, and qualitative methods and the corresponding techniques have been captured.

Level of analysis: With this component of the analysis framework, we aim at understanding whether archetype theory is more suited for any particular level of analysis. We distinguish between department (e.g., financial department), organization (e.g., a particular company), industry (e.g., law advisors), and sector (e.g., professional service firms) levels of analysis.

Complementary theory: If archetype theory is used in combination with any other theory, such observations are noted down in this component of the analysis framework. This helps us understand relevant theories that can be used as complementary to archetype theory.

After developing the analysis framework, we coded the extracted articles based on the analysis framework[3].

---

[3] The summary of the coding is available for download under http://bit.ly/2f4cJPn

4 Results

The analysis of the extant literature resulted in identifying 16 articles in which archetype theory has been employed [27-29, 31, 43-54][4]. This section presents the major findings of our review in line with the previously introduced analysis framework.

As shown in **Table 2**, the reviewed articles proposed different archetypes. However, some pairs of comparable archetypes (e.g., bureaucratic and managerial archetypes) are frequently used in different articles and only about half of the articles proposed new or substantially modified existing archetypes. Therefore, the identified archetypes are representative for similar situations and can be re-used in other studies. Further, according to the use of the theory, 13 of the articles are concerned with the analysis of change. The latter is an indication that although identification of archetypes is a considerable contribution per se, they are mainly used to better understand the dynamics of change in organizations. We have also observed that at least two different archetypes are identified in each of the reviewed articles. This observation is related to the fact that change is mainly associated with a movement between archetypes.

Table 2. Purpose of theory use and employed research methods in the reviewed articles

Category			**Reference**	**# of articles**[5]
Purpose of theory use	*Exploration of Archetypes*	Managed Professional Partnership / Managed Professional Business	[43, 44]	2
		Bureaucratic / Managerial	[45-47]	3
		Kitchen Table Boardroom / Executive Office	[29, 49, 50]	3
		Others	[27, 28, 31, 48, 51-54]	8
	Analysis of change		[31, 43-54]	13
Research Method	*Conceptual*	Literature Analysis	[27-29]	3
	Qualitative	Semi-Structured Interviews	[31, 45-48, 50, 51, 53, 54]	9
		Secondary Source Analysis	[31, 45-48, 50, 52, 53]	8
		Meeting Observation	[48, 54]	2
	Quantitative		[43, 44]	2

4.1 Organizational Archetypes

By comparing the archetypes investigated in the reviewed articles, we were able to identify patterns of how archetypes are typically described. It is out of scope of this

[4] Literature analysis uncovers 5 *seminal* articles of archetype theory [35, 38, 39, 55, 56], which are used to develop our analysis framework. The derived analysis framework is used to analyze 16 *application* articles.

[5] The same article may be assigned to multiple categories. For instance, Haki & Legner [31] applied qualitative research through semi-structured interviews and secondary source analysis.

article to describe each particular archetype in detail[6]. We rather aimed at explaining how archetypes have been described in the reviewed articles to eventually synthesize how archetype theory can actually be applied. In line with the analysis framework, we distinguish between the structural arrangement and the interpretive scheme of archetypes.

Structural Arrangement: In order to describe the structural arrangement, the work of Cooper et al. [55] is often a starting point in the reviewed articles [44, 46, 54, 55]. Cooper et al. posit that structural arrangement can be defined through both *structures* and *systems*. *Structures* are generally classified by considering the degree of differentiation and integration. Differentiation has been identified through, for instance, considering the level of specialization amongst teams [28] or through the range of different disciplines within organizational entities [48]. Integration is considered to be expressed by the location of the decision power [49], the information flow [55], and the degree of commonly applied rules and procedures. *Systems*, here mainly to be understood as processes, are generally classified into strategic control, marketing control, financial control, and operating control. Reviewed articles propose to measure strategic control through, for example, the degree of strategic freedom of different organizational entities [43, 44]. Marketing and financial control can be observed by the tolerance regarding financial and marking targets [43], compensation systems, and systems for performance appraisal [46]. Operating control finally may be expressed by the degree of centralization of control and information systems [46].

Interpretative scheme: Less consistent are the descriptions used for the interpretative scheme. This is due to the fact that intangibles (such as values and beliefs) are much more difficult to define and measure than tangibles (such as an organization's structure). For instance, in order to distinguish between "partnership" and "managed" archetypes, authors looked at the perceived purpose of an organization, also described as an organization's "raison d'être" [54]. For partnership, this is often the exchange of knowledge with peers, whereas for managed organizations this would be the increase of productivity [43, 44]. As another example, for sport organizations, differentiation has been made between their degree of professionalism in terms of their target definition (e.g., sport as a leisure activity vs. sport as a profession) [29]. Other scholars also differentiated different interpretive schemes along the underlying principles when taking decisions, for example regulation-oriented versus efficiency-oriented decision making principles [45, 46].

4.2 Change

Since the concept of change tracks is an integral part of archetype theory, prior to our review we expected that the articles to take up and refer to this concept [38]. Surprisingly, we identified only few articles [49, 54, 55] that explicitly distinguish between change tracks as explained in Section 2. Instead, Liguori [47], for example, employs a rather basic construct of change tracks and distinguishes between incremental and radical changes. Incremental changes are considered as modifications

[6] Brock [57] provides detail specifications of the identified archetypes in the literature.

of the structural arrangements only, whereas radical changes also involve changes in the interpretative scheme. This conception is confirmed by the other authors, highlighting the crucial role of the interpretative scheme in change processes [44, 46]. It is noteworthy that Kirkpatrick & Ackroyd [52] demonstrate, how change does not necessarily lead to a new archetype but may also be reflected in an adjustment within the current archetype. Therefore, change can be examined through both adjustment within an archetype and movement between different archetypes.

Further, instead of focusing on the change process as such, authors were interested in understanding why and how a change process is initiated. Authors understand change of the archetype as a reaction to environmental and contextual pressures, which are filtered by organizations through an internal process of interpretation and attribution of meanings [47, 52]. Frequently given examples for such pressures are globalization [28, 44, 52], (de-)regulation/change in government policy [28, 44, 52], change in client needs [44], technological progress [52], as well as capacity for action in terms of both technical and managerial/leadership capabilities [47].

4.3 Research Method, Level of Analysis, and Complementary Theories

In this section we describe the employed research methods in the reviewed articles followed by a discussion of the level of analysis and complementary theories. The majority of the reviewed articles employed *qualitative* research methods and case study research in particular (see **Table 2**).

In case studies, semi-structured interviews, and the review of secondary sources, such as documents, reports, presentations, and media articles were common to identify archetypes and/or changes among archetypes. Indicators for the structural arrangement were, for example, the degree of integration expressed by the decision power of the headquarter [28]. The interpretative scheme was, for example, assessed by capturing the underlying principles during the decision making process [46, 47]. In addition, two articles identified the observation of meetings as an appropriate technique to extract values and decision making processes [48, 54].

Change was identified in two different ways: Either by comparing different cases at a single point in time, or by carrying out longitudinal case studies. The decision for either of the two options is thereby depending on the underlying research question. Liguori [47], for example, was interested in why similar organizations react differently to the same kind of change, whereas Carter and Mueller [48] were interested in the change process of one organization between two archetypes. This implies that whenever the dynamics of change of one particular organization are of interest, longitudinal studies are more appropriate.

Conceptual research was only conducted in articles focusing on rather abstract research topics such as ideal types of governance [27] or synthesis on archetypes described by the other authors [29].

Concerning *quantitative* methods, only two articles applied archetype theory through a quantitative research design [43, 44]. Both articles aimed at examining change of archetypes in different groups of professions (architects and law firms). To this end, they did so by taking two already defined archetypes and assigned the organizations to one of them based on the answers received in a questionnaire.

The reviewed articles investigated archetypes and change, on different *levels of analysis*. Three articles focused on the department level, for instance, the departments of local government and change in the accounting system [45-47]. Two articles focused on single companies and were classified as research on the level of organization [31, 48]. The majority of the reviewed articles examined archetypes and change on the industry level. Two conceptual articles also focused on the sector of legal advisors and auditors [27, 28].

Besides the employed research methods and the level of analysis, we were also interested in identifying *theories* that were used complementary to archetype theory. We consider configuration theory (more precisely, the typology and taxonomy aspects) as well as agency theory to be relevant, because they were used in the reviewed articles to facilitate exploration of archetypes and explanation of change in archetypes. Harlacher und Reihlen [27] employed configuration theory to identify governance taxonomies and compared them with existing archetypes in the literature. Pinnington [44] employed agency theory to better explain change in archetypes. The use of agency theory helped Pinnington [44] explain changes in the decision making system (structural arrangement) for cases, where the ownership of organizations has changed [44]. According to agency theory, control mechanisms are intensified in cases, where the ownership and the management of the same organization are separated (e.g., shareholder vs manager). This may lead to a change in archetypes because not only the structures but also the values are changed.

5 Discussion

This article starts with the premise that the study of configurations should account for the underlying change mechanisms that bring about the creation or emergence of the respective configurations. As such, the simultaneous study of change and configurations results in profound insights on the dynamics of configurations, their development in a series of change events, and eventually gives meaning to their aspects and specifications.

IS scholars have been striving to explain change and proposed a variety of approaches to identify optimal configurations. To this end, various theoretical lenses have been employed to study change (e.g., evolutionary/Darwinian approach) and optimal configurations (e.g., contingency theory). Relying on the constitutive constructs and theoretical premises of archetype theory, this theory provides a theoretically sound basis to not only explain changes but also to explore configurations. Therefore, it can be employed in studying various IS phenomena in which identifying configurations and explaining the underlying change mechanisms are central. To elaborate the employment of the archetype theory and its contributions in prospective IS research, we provide two exemplary implications on both research streams namely, optimal configurations and change:

The study of optimal configurations is quite dominant in, for instance, IT governance. As pointed out in a literature review on IT governance by Brown and Grant [58], the extant literature has been dominated by either configurations of IT governance [7, 59, 60] — introducing centralized, decentralized, and federal governance modes — or by contingency factors influencing IT governance structure [8]. Nevertheless, existing studies mainly prescribe optimal IT governance structures while the underlying mechanisms that bring about these structures remain obscure. Therefore, the promoted lens of archetype theory proposes a dynamic, non-deterministic approach to explain how and why different IT governance structures arise. It also brings up the possibility of establishing different IT governance structures under the same organizational contingencies and further explains how this non-deterministic process is as such. Moreover, concerning the change stream of research, extant research gives rise to the nature of IS change so that change is not solely or even mainly incremental and cumulative, but rather is episodic and punctuated [61, 62]. These studies mainly lay emphasis on explaining change mechanisms while the emerged configurations form this dynamic process is of outmost interest for different IS phenomena. As such, the explored relation between change processes and the emerged configurations can not only give rigorous meaning to the dynamics of configurations and their occurrence, but also explore typological configurations that can be observed in different contexts and situations. Therefore, the use of archetype theory can systematically guide prospective IS research to not only explain the nature of change but also explore configurations.

To apply archetype theory in IS research, our review motivates longitudinal case studies to rigorously reflect dynamics of configurations over time. We encourage mixed-methods research to not only explore configuration but also to confirm their generalizability in a larger extent. Our review confirms applicability of archetype theory on different levels of analysis. Therefore, we not only see value in applying archetype theory on a more granular level, where local variants of organizational specialties may be considered, but also in studies that concern multiple levels of analysis. Finally, our review reveals that, owing to its explanatory power, the theoretical constructs of archetype theory can be synthesized with other theories (e.g., agency theory) to be used as complementary lenses in providing thorough explanations of IS phenomena.

Our literature review is limited to publications in scholarly journals and AIS conference proceedings. This is a limitation in terms of coverage of relevant research (e.g., conference proceedings in other disciplines). However, this restriction ensures a certain quality level of the reviewed articles while including conference-level contributions from IS. Our study reveals that archetype theory has not gained much attention in IS literature (only one article, [31]). Therefore, our review provides insights on how this theory can be applied in prospective IS research.

6 Conclusion

While spotlighting the necessity of simultaneous investigation of configurations and change, this study contributes to the existing body of IS knowledge through introducing archetype theory as a new theoretical lens, which guides IS scholars in such investigations. Through a literature review, this study carefully extracts theoretical constructs of archetype theory and investigates why, how, with which approaches, and through which methods archetype theory has been applied in the extant literature.

The concept of archetype provides a basis to systematically describe configurations in terms of their structural arrangement as well as the values and beliefs that reinforce these arrangements. The theoretical assumptions on change in archetype theory also provides a rich explanation on the dynamics of configurations. As such, this theory helps IS scholars elaborate on change not only through demonstrating movements among different archetypes, but also through indicating adjustments within an archetype.

References

1. Doty, D.H., Glick, W.H., Huber, G.P.: Fit, equifinality, and organizational effectiveness: A test of two configurational theories. Academy of Management Journal 36, 1196-1250 (1993)
2. Miller, D.: Configurations of Strategy and Structure: Towards a Synthesis. Strategic Management Journal 7, 233-249 (1986)
3. Short, J.C., Payne, G.T., Ketchen, D.J.: Research on organizational configurations: Past accomplishments and future challenges. Journal of Management 34, 1053 -1079 (2008)
4. Xue, Y., Liang, H., Boulton, W.R.: Information technology governance in information technology investment decision processes: The impact of investment characteristics, external environment, and internal context. MIS Quarterly 31, 67-96 (2008)
5. Nickerson, R.C., Varshney, U., Muntermann, J.: A method for taxonomy development and its application in information systems. European Journal Of Information Systems 22, 336-359 (2013)
6. El Sawy, O.A., Malhotra, A., Park, Y., Pavlou, P.A.: Research commentary-seeking the configurations of digital ecodynamics: It takes three to tango. Information Systems Research 21, 835-848 (2010)
7. Brown, C.V.: Examining the Emergence of Hybrid IS Governance Solutions: Evidence from a Single Case Site. Information Systems Research 8, 69-94 (1997)
8. Sambamurthy, V., Zmud, R.W.: Arrangements for information technology governance: A theory of multiple contingencies. MIS Quarterly 23, 261-290 (1999)
9. Madapusi, A., D'Souza, D.: Aligning ERP systems with international strategies. Information Systems Management 22, 7-17 (2005)
10. Markus, M.L., Tanis, C., Van Fenema, P.C.: Enterprise resource planning: multisite ERP implementations. Communications of the ACM 43, 42-46 (2000)
11. Lyytinen, K., Damsgaard, J.: Inter-organizational information systems adoption–a configuration analysis approach. European Journal Of Information Systems 20, 496-509 (2011)
12. Tsoukas, H., Chia, R.: On organizational becoming: Rethinking organizational change. Organization Science 13, 567-582 (2002)
13. Sabherwal, R., Robey, D.: An empirical taxonomy of implementation processes based on sequences of events in information system development. Organization Science 4, 548-576 (1993)
14. Markus, M.L., Robey, D.: Information technology and organizational change: causal structure in theory and research. Management Science 34, 583-598 (1988)

15. Lyytinen, K., Newman, M.: Explaining information systems change: a punctuated socio-technical change model. European Journal of Information Systems 17, 589-613 (2008)
16. Luna-Reyes, L.F., Zhang, J., Gil-García, J.R., Cresswell, A.M.: Information systems development as emergent socio-technical change: a practice approach. European Journal of Information Systems 14, 93-105 (2005)
17. Orlikowski, W.J.: Improvising Organizational Transformation Over Time: A Situated Change Perspective. Information Systems Research 7, 63-92 (1996)
18. Bakos, J.Y., Kemerer, C.F.: Recent applications of economic theory in information technology research. Decision Support Systems 8, 365-386 (1992)
19. Orlikowski, W.J., Barley, S.R.: Technology and institutions: What can research on information technology and research on organizations learn from each other? MIS Quarterly 25, 145-165 (2001)
20. Walsham, G.: Doing interpretive research. European Journal of Information Systems 15, 320-330 (2006)
21. Straub, D.: Editorial: Does MIS have native theories. MIS Quarterly 36, iii-xii (2012)
22. Truex, D., Holmström, J., Keil, M.: Theorizing in information systems research: A reflexive analysis of the adaptation of theory in information systems research. Journal of the Association for Information Systems 7, 797-821 (2006)
23. Gregor, S.: Theory–still king but needing a revolution. Journal of Information Technology 29, 337-340 (2014)
24. Mignerat, M., Rivard, S.: Positioning the institutional perspective in information systems research. Journal Of Information Technology 24, 369-391 (2009)
25. Nielsen, J., Mathiassen, L., Newell, S.: Theorization and translation in information technology institutionalization: evidence from Danish home care. MIS Quarterly 38, 165-186 (2014)
26. Guillemette, M.G., Paré, G.: Toward a New Theory of the Contribution of the IT Function in Organizations. MIS Quarterly 36, 529-551 (2012)
27. Harlacher, D., Reihlen, M.: Governance of Professional Service Firms: A Configuration Approach. Business Research 7, 125-160 (2014)
28. Brock, D.M., Powell, M.J., Hinings, C.R.: Archetypal Change and the Professional Service Firm. Research In Organizational Change And Development 16, 221-251 (2007)
29. Kikulis, L.M., Slack, T., Hinings, B.: Institutionally Specific Design Archetypes: A Framework for Understanding Change in National Sport Organizations. International Review For The Sociology Of Sport 27, 343-368 (1992)
30. Fiss, P.C.: A Set-Theoretic Approach to Organizational Configurations. Academy Of Management Review 32, 1180-1198 (2007)
31. Haki, M.K., Legner, C.: The dynamics of IS adaptation in multinational corporations: a new theoretical lens. In: 2013 International Conference on Information Systems (ICIS 2013). Association for Information Systems (2013)
32. Van de Ven, A.H., Ganco, M., Hinings, C.R.: Returning to the Frontier of Contingency Theory of Organizational and Institutional Designs. Academy of Management Annals 7, 393-440 (2013)
33. Rackoff, N., Wiseman, C., Ullrich, W.A.: Information Systems For Competitive Advantage: Implementation of a Planning Process. MIS Quarterly 9, 285-294 (1985)
34. Powell, W.W.: The New Institutionalism. In: Clegg, S., Bailey, J.R. (eds.) International encyclopedia of organization studies. Sage Publications (2007)
35. Greenwood, R., Hinings, C.R.: Understanding Strategic Change: The Contribution of Archetypes. Academy of Managemenl Journal 36, 1052-1081 (1993)

36. Meyer, A.D., Tsui, A.S., Hinings, C.R.: Configurational Approaches to Organizational Analysis. The Academy of Management Journal 36, 1175-1195 (1993)
37. Greenwood, R., Oliver, C., Suddaby, R., Sahlin-Andersson, K. (eds.): The SAGE Handbook of Organizational Institutionalism. Sage Publications, London (2008)
38. Greenwood, R., Hinings, C.R.: Organizational Design Types, Tracks and the Dynamics of Strategic Change. Organization Studies 9, 293–316 (1988)
39. Laughlin, R.C.: Environmental Disturbances and Organizational Transitions and Transformations: Some Alternative Models. Organization Studies 12, 209–232 (1991)
40. Webster, J., Watson, R.T.: Analyzing the Past to prepare for the Future: Writing a Literature Review. MIS Quarterly 26, 13-23 (2002)
41. Miller, D.: Towards a New Contingency Approach: The Search for Organizational Gestalts. Journal Of Management Studies 18, 1-26 (1981)
42. Fettke, P.: State-of-the-Art des State-of-the-Art. Eine Untersuchung der Forschungsmethode „Review" innerhalb der Wirtschaftsinformatik. Wirtschaftsinformatik 48, 257–266 (2006)
43. Pinnington, A., Morris, T.: Archetype Change in Professional Organizations: Survey Evidence from Large Law Firms. British Journal Of Management 14, 85-99 (2003)
44. Pinnington, A., Morris, T.: Transforming the Architect: Ownership form and Archetype Change. Organization Studies 23, 189-210 (2002)
45. Liguori, M., Steccolini, I.: Accounting change: explaining the outcomes, interpreting the process. Accounting, Auditing & Accountability Journal 25, 27-70 (2011)
46. Liguori, M.: The Supremacy of the Sequence: Key Elements and Dimensions in the Process of Change. Organization Studies 33, 507-539 (2012)
47. Liguori, M.: Radical Change, Accounting and Public Sector Reforms: A Comparison of Italian and Canadian Municipalities. Financial Accountability & Management 28, 437-463 (2012)
48. Carter, C., Mueller, F.: The 'long march' of the management modernizers: Ritual, rhetoric and rationality. Human Relations 55, 1325-1354 (2002)
49. Kikulis, L.M., Slack, T., Hinings, C.R.: Sector-Specific Patterns of Organizational Design Change. Journal Of Management Studies 31, 67-100 (1995)
50. Hinings, C.R., Thiabault, L., Slack, T., Kikulis L. M: Values and Organizational Structure. Human Relations 49, 885-916 (1996)
51. Stensaker, B., Frølich, N., Huisman, J., Waagene, E., Scordato, L., Pimentel Bótas, P.: Factors affecting strategic change in higher education. Journal of Strategy and Management 7, 193-207 (2014)
52. Kirkpatrick, I., Ackroyd, S.: Transforming the professional archetype? The new managerialism in UK social services. Public Management Review 5, 511-531 (2003)
53. Dent, M., Howorth, C., Mueller, F., Preuschoft, C.: Archetype transition in the German health service? The attempted modernization of hospitals in a North German state. Public Administration 82, 727-742 (2004)
54. Stanley-Clarke, N., Sanders, J., Munford, R.: Implementing a new governance model: Lessons from a New Zealand statutory mental health organisation. Journal of Health Organization and Management 30, 494-508 (2016)
55. Cooper, D.J., Hinings, B., Greenwood, R., Brown, J.L.: Sedimentation and Transformation in Organizational Change: The Case of Canadian Law Firms. Organization Studies 17, 623-647 (1996)
56. Greenwood, R., Hinings, C.R.: Understanding Radical Organizational Change: Bringing together the Old and the New Institutionalism. The Academy of Management Review 21, 1022-1054 (1996)

57. Brock, D.M.: The changing professional organization: A review of competing archetypes. International Journal of Management Reviews 8, 157-174 (2006)
58. Brown, A.E., Grant, G.G.: Framing the Frameworks: A Review of IT Governance Research. Communications of the Association for Information Systems 15, 696-712 (2005)
59. Weill, P., Ross, J.W.: IT Governance - How Top Performers Manage IT. Harvard Business School Press, Boston (2004)
60. Brown, C.V., Magill, S.L.: Alignment of the IS Functions with the Enterprise: Toward a Model of Antecedents. MIS Quarterly 18, 371-403 (1994)
61. Lyytinen, K., Newman, M.: Explaining information systems change: a punctuated socio-technical change model. European Journal of Information Systems 17, 589-613 (2008)
62. Sabherwal, R., Hirschheim, R., Goles, T.: The Dynamics of Alignment: Insights from a Punctuated Equilibrium Model. Organization Science 12, 179-197 (2001)

Testing Technical Feasibility in CPS Development Projects

Moritz von Hoffen[1], Friedrich Chasin[1], Martin Matzner[2], Florian Plenter[1], and Jan H. Betzing[1]

[1] University of Münster – ERCIS, Leonardo-Campus 3, 48149 Münster, Germany,
{moritz.von.hoffen,friedrich.chasin,
florian.plenter,jan.betzing}@uni-muenster.ercis.de
[2] University of Erlangen-Nürnberg, Digital Industrial Service Systems, Nuremberg, Germany
martin.matzner@fau.de

Abstract. Cyber-physical systems (CPSs) are service systems that connect a product's physical and computational elements through telecommunication networks. Typically, the processes in CPSs are executed on this physical and computational infrastructure. As the developing of new CPS is costly, testing and validating a CPS's design at an early stage of development is desirable in order to avoid potential bad investments. The high development and potentially high hardware costs, however, make it difficult to create a full CPS prototype only for testing. This work uses Trkman's critical success factors of business process management (BPM) as a theoretical lens and identifies "technical-feasibility fit" as an additional complementary success factor. Based on these factors, we develop a method for creating CPS testbeds that allow testing of CPSs at lower costs at an early stage of the development. We demonstrate the method's application by a case in which we develop a testbed for an electric vehicle charging service.

Keywords: Cyber-physical Systems, Critical Success Factors, Prototyping

13th International Conference on Wirtschaftsinformatik,
Februrary 12-15, 2017, St. Gallen, Switzerland

von Hoffen, M.; Chasin, F.; Matzner, M.; Plenter, F.; Betzing, J. H. (2017): Testing Technical Feasibility in CPS Development Projects, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 589-603

1 Introduction

Service is the value created in relational interaction processes [1] that connect a company to several "collaborators" [2, p. 492] such as partners, employees, and suppliers. All entities together form the *service system*, which VARGO and LUSCH defined as a dynamic configuration of four types of resources, i.e., people, technologies, organization, and information [3]. *Service Development* refers to a firm's approach to creating new service offerings and has been described as a cyclical process that includes various planning and implementation activities at the progressive stages of "Design", "Analysis", "Development", and "Launch" [4]. Notably, a service development project can also return to earlier stages if later planning and development activities require modifications of the service concept. Against this background, service development requires that service concepts are tested repeatedly for their business value and for their operational feasibility [5]. For instance, shortly after initial idea generation, firms typically evaluate the service concept ideas through a "screening" [6]. However, the individual development stage activities related to the design of the service, processes, and the actual system require more exhaustive testing [4]. Business value embodied in, e.g., profitability, growth and reward potential, as well as competitive advantage [6, 7] is typically assessed using conventional qualitative and quantitative market research techniques like surveys, focus groups, one-on-one interviews and conjoint analysis [8].

In contrast, the testing of the operational feasibility of a service concept requires to look deeply into the service system's value creation processes as well as the technological and informational resources they use. "Prototyping" is one approach to achieve rapid customer-centric service experimentation [9]. In this context, it is an important question how *service prototyping* can be used to "materialize an integrated set of service system components, such as the people, the process, the technology, and the physical evidence" [10, p. 137]. According to OSTROM et al., prototyping has not received sufficient attention in service research, and thus, they feature service prototyping in their recent list of important service research areas [10].

Especially in the domain of cyber-physical systems (CPSs), technology components are of particular significance. Typically, CPSs connect (remote) computational and physical entities, e.g., sensors and actuators, via global computational networks [11]. In this context, *prototyping* is highly important for this type of service system because of the requirement that complex technical infrastructures have to be built at early stages of the service development—even before the progress and the processes can be fully tested.

Against this backdrop, the present paper addresses the following research question: *How can service prototyping materialize the process and technology components of cyber-physical systems?* The contribution of this paper lies in the design of a method for creating CPS testbeds. We intend to improve CPS service development by facilitating prototyping and testing for operational feasibility at early stages of the development process and at reasonable costs.

The development of the method is informed through the theory of *task-technology fit (TTF),* a theoretical lens that has been applied previously in business process management (BPM) [12, 13]. The TTF theory helps to assess whether certain technologies are appropriate to a given process. Therefore, this paper also seeks to synthesize research on prototyping in service development and research on success factors in BPM, which have so far been considered only separately.

The remainder of this article is structured as follows: The next section gives background on the testing of CPS with regard to the process perspective and success factors. Then we explain our research approach, followed by the method and a demonstration of its application in a project that develops a service for electric vehicle (EV) charging. The evaluation section provides first evidence of the method's usefulness. A discussion of our results follows, and the final section concludes the article.

2 Research Background

2.1 Challenges in Testing CPS

CPSs are specific service systems including networked computational systems that are partly embedded into physical objects [11]. Sensors and actuators connect the physical and digital worlds. An ever-growing number of CPSs, which have become ubiquitous in every-day life, generate a vast amount of data, with typical applications ranging from smart grids [14], physical infrastructures in transportation [15], traffic and process control to automotive and medical systems [16].

CPSs are complex systems with complex processes that typically run on expensive hardware. In particular, the embedding of physical components requires higher standards for reliability and safety as system failures can result in severe damages, e.g., of the environment [17]. Embedded systems such as driving assistance or brake control are examples of CPSs integrated in every-day systems, whose failures can result in serious consequences for the public. Moreover, the behavior of CPSs cannot always be predicted.

While traditional end-to-end business processes are implemented within or across a few application systems, processes in CPSs add another layer of complexity. In effect, parts of the business logic are shifted into these embedded systems [18]. From the business perspective, addressing the challenges posed by the nature of CPSs requires considerable investment. Failing in the latter stages of the development due to miss-specified processes that are unable to execute within and across the CPS can be costly, so guarding against such situations is critical for managing and executing business processes.

2.2 Process-Focus in Service Design and Testing

Testing CPS for operational feasibility is of great importance throughout the various stages of service development [5]. Prototyping has been discussed as a promising ap-

-proach to achieve a balance between receiving early insight on the feasibility and the costs associated with the testing activities [10]. The key intellectual challenge in service prototyping is to achieve an integrated service experimentation that materializes all relevant components of the service systems in a way so that the service stakeholder can make sense of the service and make reasonable decisions about the progress of the development project [9]. The scope of this paper has been set to processes and technology, which are the most significant components of CPS service systems.

The BPM literature has put forth constructs, models, and theories that help to study the relationship between processes and technology. Notably, in an attempt to identify critical success factors (CSFs) for BPM, TRKMAN demands "continuous improvement efforts" for BPM and two types of "fit" for business processes [12, p. 126]. The "fit between business environment and business processes" has been explained by the *contingency theory* [19], which in essence states that there exists no universal or "best way" to manage an organization. Instead, achieving an appropriate organization is contingent to various internal and external constraints. Accordingly, business processes have to be designed so that they meet the constraints of the process environment. The "fit between business processes and technology" [12, p. 127] has been explained through *task-technology fit (TTF)*—a theory that identifies that a positive impact of information technology (IT) investments on organizational performance is subject to matching IT and business processes.

The need for "dynamic improvement" of business processes is justified by the theory of *dynamic capabilities*, which postulates that organizations need to address changing environments through the ability to integrate, build, and reconfigure internal and external competences. Therefore, business processes need to be reviewed for both types of fit continuously.

We focus on TTF, which provides means to study the CPS's *process* and *technology* service components in conjunction. While testing in software development projects already accounts for about one third of development cost [20], the testing and validation of the distributed and embedded components of a CPS is even more complex and costly [21] and thus underlines the importance of a proper TTF.

3 Research Approach

To approach the problem of testing the operational feasibility of a service concept in the context of a CPS throughout different stages of the service development process, we perform two research activities: At first, we aim at the derivation of a framework for critical success factors in BPM from the extant literature. This step is required to examine and categorize different state-of-the-art CSFs to identify the gap and motivate the extension of the framework with an additional CSF of testing the operational feasibility of the service concept. The second strand of our research deals with the development of a method that is capable of closing the identified gap.

Table 1. Research Steps

#	Step	Activities	Outcomes
1	*Analyze CSFs*	Investigate extant CSFs Align CSFs and BPM Analyze applicability in use case scenario	Framework Identified gap
2	*Propose Extension*	Propose testbedding as an extension to the TTF	Motivation Proposed extension
3	*Method Design*	Design the method Apecify activities	Method instantiation
4	*Demonstration*	Demonstrate the extension Iterative execution Adopt continuous improvement	Evaluation

Table 1 summarizes the steps undertaken in this research and the generated outcomes. In the first step, we identify relevant CSFs in the extant literature (see CSFs labelled with [*] in Table 2). Taking the specific requirements imposed on service testing in the CPS context, we propose the extension of the general framework of success factors in BPM with an additional CSF (see CSF_{2+} in Table 2). We then design the method suitable for ensuring and warranting that the chosen technology set for the service at hand aligns with the corresponding business processes. Finally, we complement the demonstration of the method with a discussion of the proposed approach against related testing procedures.

Table 2. Overview of *Critical Success Factors* ([*] according to [12])

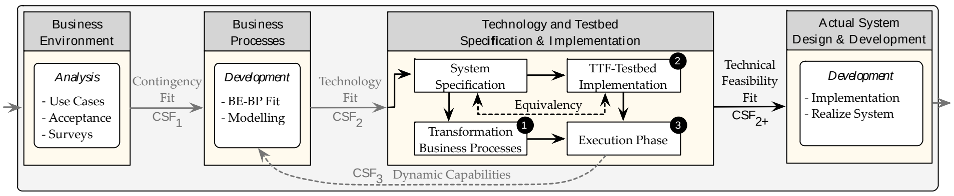
CSF_1[$\star$]: *Contingency Fit*		
Activities	***Outcomes***	***Literature***
•Evaluation of business engineering •Alignment of strategy with business engineering •Conceptualization and modeling of business processes	Contingency of the business engineering and business processes	[12, 13, 19, 22]
CSF_2[$\star$]: *Task-Technology Fit*		
Activities	***Outcomes***	***Literature***
•Alignment of supporting IT with business processes	Technological compatibility of IT and business processes	[12]
CSF_{2+}: *Technical Feasibility Fit*		
Activities	***Outcomes***	***Literature***
•Mapping of business processes to a state-based representation •Identify the technologies required for testbed implementation •Assemble the testbed •Execution of business processes in testbed	Feasibility of selected technological components	—
CSF_3[$\star$]: *Dynamic Capabilities*		
Activities	***Outcomes***	***Literature***
•Continuous improvement of business processes •Responding to changes in the business engineering	Mature business processes achieved through continuous improvement	[12, 19]

4 A Method for Creating CPS Testbeds

Against the backdrop of high risks associated with business processes relying on CPSs, we enhance TRKMAN's model of CSFs [12] by an additional CSF that follows the TTF. We introduce the specification and implementation of a testbed as means of ensuring the *technical feasibility fit* between the chosen technology set and the business processes (cf. Figure 1). In effect, a testbed combines virtual, simulated, and physical components into a configurable experimental setup for testing [23]. In an ideal world, the behavior and properties of the testbed are equivalent to ones of the specified service system. Thus, we use the term *testbed equivalency*.

Our work aims at achieving an optimal TTF with a technology ensemble feasible to execute the business processes, while treating the remaining activities required to address further CSFs as a *black box*. Assessing the TTF can be expensive, which is especially true for distributed processes that run on heterogeneous and specialized hardware. Prototype development with a testbed combines the benefits of early testing and validation with cost savings, because the actual hardware roll-out can be postponed until the testbed has been used to validate the correct execution of all involved (business) processes. Hence, prototyping can "reduce the chances of costly new service failures" [10, p. 137].

Testbeds have to correctly imitate the execution of the business processes, and thus, require a precise specification of the target system's behavior. We therefore limit the scope to business processes that use standardized and established technologies, techniques and protocols, so that the behavior of the system can be anticipated.



4.1 Steps of the Method

Figure 1. Embedding of Testbedding into the Framework of CSFs for BPM

Figure 1 locates the proposed *technology and testbed specification and implementation* within the process and framework of CSFs in BPM. The activity blends in after CSF_1 and CSF_2 have been achieved through the contingency and technology fits. At this point, the business processes have been modelled and formalized. Based on the business processes and underlying standards, a set of technologies, i.e. software and hardware components, has been chosen. The testbed method consists of three steps: First, the resulting business processes must be *transformed (1)* into a state-based representation. Simultaneously, the *testbed implementation (2)* is performed. The testbed equivalency to the actual system must be assured through the equivalency of the testbed specification to the system specification. Subsequently, the implemented and configured testbed is put to use in the *execution phase* (3) by executing the transformed business processes from step (2).

(1) Transformation of Business Processes into a State-based Representation: Business processes are typically represented as models and the Event-driven Process Chain (EPC) and Business Process Model and Notation (BPMN) are arguably the most prominent graphical process modeling languages in both, academia and practice. To ensure correct process execution and *soundness*, one needs to transform the EPC or BPMN models into a *state-based* representation. The utilization of a state-based representation allows to precisely comprehend the current state of the process and check every state transition for compliance. In this context, *Petri nets* are recommendable [24] which is justified by the large existing body of knowledge on formal validation of Petri nets [25]. Moreover, for the actual transformation, one can make use of existing and well-tried methods for *model-to-model transformation* to convert the business processes into Petri nets. The mapping itself is straightforward: "*tasks* are modeled by *transitions*, *conditions* are modeled by *places*, and *cases* are modeled by *tokens*." [26, p. 15]. A *marking* can be understood as a snapshot that reflects the Petri net's state at a certain point in time. In order to make the state transitions of the resulting Petri net transparent, the individual states must be represented in such a way that they are *observable*. This approach is similar to *lean manufacturing* or *Andon systems* where certain situations are signaled. Several options are conceivable for providing such an output like displays, acoustic signals, and light-emitting diodes (LEDs). Once a suitable and observable state representation has been decided upon (e.g., LEDs), a coherent mapping of the individual Petri net markings, i.e., states, must be developed. A naïve solution is to assign an individual LED to each place on the Petri net, which would visualize the presence of a token at the corresponding place. However, the number of places in the resulting Petri net can be large for complex business processes. This can be mitigated by using different states of the same signal emitter to code the marking, e.g., using multi-colored LEDs and modes like *on/off* or *blinking/pulsing* in different intervals. A display can also be attached that can be used to output the state as well as accompanying information such as enabled transitions or a history of states, which can be used for *backtracking* purposes to achieve full coverage of the process.

(2) Testbed Implementation: Based on the results of the TTF assessment, a testbed has to be specified that ensures *equivalency* to the technology set intended for the implementation of the productive IT infrastructure. Due to heterogeneity in required capabilities among various use cases and domains, the hardware selection process needs to be considered individually. However, a careful evaluation of the underlying task-technology fit is mandatory to provide a tangible basis for choosing suitable hardware components for the intended testbed. Naturally, the selected hardware should be capable to imitate the actual productive component of the CPS.

Important activities in this development stage include the assessment and comparison of different hardware and software vendors and sources, as well as their compatibility. Active support and maintenance should be taken into account as well. The decision regarding a suitable means for the output depends on many factors like the total number of devices in the testbed, the processes' complexity, and the degree of concurrency. Depending on which output mean(s) is/are chosen, a formal mapping and representation of the different states must be defined. Finally, the testbed device(s) is/are assembled and programmed so that it can emulate the intended business processes.

(3) Execution Phase: The set of business processes to be tested should be compiled beforehand and in accordance with the testbed specification. The resulting *test suite* is then processed by executing the different processes in the testbed environment. In the spirit of *continuous improvement*, the processes are executed iteratively within the testbed. If an abnormality is experienced during the execution phase, this information is recorded and re-evaluated in an iterative manner in a subsequent execution round. Each execution of the processes in the testbed environment provides feedback to the specification phase until the result meets the acceptance criteria. In some cases, the testbed might also prove that a given business process is unfeasible for real-world execution. This information flow and the subsequent addressing of defects results in a demonstrated *technical feasibility* of the technology—given the assumption that the testbed and the final system are equivalent when the processes are emulated. This additional step that contemplates the initial TTF constitutes an additional CSF: *Technical Feasibility Fit*. Once the "sweet spot" in terms of robustness has been reached, the replacement of the testbed with the actual production system is approached.

5 Demonstration

5.1 Project Setting: EV Charging Infrastructure

We applied the testbed method in the domain of EV charging. EVs are charged using charging points, which combine electrical and computational components. The CPS at hand comprises a charging infrastructure of networked charging points and an information system (IS) that, among other tasks, controls the individual charging points, authorizes users to unlock a charging point and charge their vehicles, and handles the billing of charging transactions.

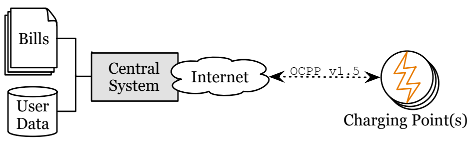


Figure 2. Central System and Charging Points

In particular, the processes for controlling the charging infrastructure, which comprises the *charging points* and a corresponding *central system* (see Figure 2), have been formalized in the *Open Charge Point Protocol (OCPP)*. The OCPP represents a *de facto standard* and protocol for the communication between the central system and the individual charging points [27]. The communication is realized by sending OCPP-based SOAP requests over HTTP. According to the OCPP, charging points can be in one of four states and different messages are used to either initiate or communicate a change of the state as visualized in Figure 3. For instance, an *expired reservation* is to be detected by the charging station itself which will then change its status to *available* whereas a request to *cancel a reservation* is sent by the central system to a specific charging station.

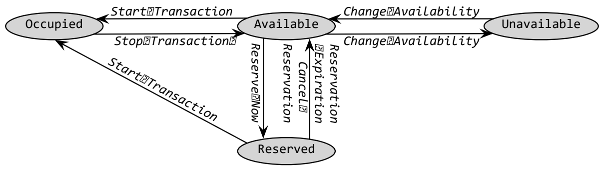


Figure 3. Transition System of a Charging Point According to the *OCPP v1.5*

5.2 Method Application

We perform the three aforementioned steps of the method (cf. Figure 1):

(1) Transformation of Business Processes into a State-based Representation: As the testbed is supposed to ensure that the business processes and the real world are compatible, a transformation into a state-based representation is performed. This is realized by transforming the individual BPMN models into a Petri net representation. In our case, we transformed the business processes that have been specified for the central system and a charging station (cf. Figure 2). To make the different states "experiencable" and observable, we relied on signaling using LED states to represent the states of the charging point:

$$\text{State} \equiv \{[\text{LED}_{\text{red}} = x_1], [\text{LED}_{\text{yellow}} = x_2], [\text{LED}_{\text{green}} = x_3]\},\ x_i \in \{\text{off}, \text{on}, \text{pulse}, \text{blink}\} \quad (1)$$

Figure 4 illustrates how the individual states of the *BPMN model* are mapped to a *Petri net* and specific *LED states*. Each marking of the resulting Petri net corresponds to a unique LED allocation.

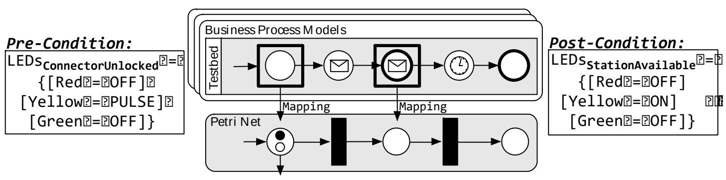


Figure 4. *Mapping* of BPMN to *Petri Net* and *LED States*

(2) Testbed Implementation: The architecture of charging stations mandates several functional requirements for the testbed implementations: reading of NFC cards, mobile

Internet connectivity, and exchange of messages with the central system based on OCPP v1.5. The publicly available Web Service Description Language (WSDL)[1] files were used for the specification and software development for the testbed and the corresponding central system. We then engineered a testbed device to resemble a charging station for the testbedding.

LEDs indicate the state of charging station as in Figure 3 (i.e., idle, reserved, out-of-business, connector (un-)locked, charging, charging finished. In addition to the LED signals, the transition sequences are logged and shown on a display. Because of its low cost, versatility, broad support, and active community, a *RaspberryPi* Model B[2] served as the basis for the development of the testbed .

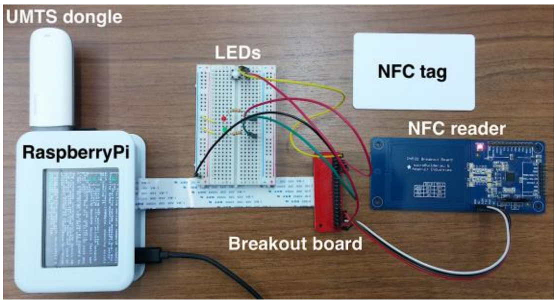


Figure 5. The *Testbed* Devices

Figure 5 shows the testbed device imitating a charging station including a near field communication (NFC) reader and the connected LEDs. The device itself is mounted in a plastic housing and also features a touchscreen display. The device connects to the Internet via universal mobile telecommunications system (UMTS) or wireless local area network (WLAN). Finally, the previously developed software was deployed on the testbed device.

(3) Testbed Execution: After specifying the mappings between the states of the business process and the states of the LEDs, the testbed is used to execute the processes. A continuous improvement cycle is included in the execution. All methods in the standard were tested against a OCPP-compliant central system. Correct business process execution is evaluated for each execution round by observing the specific outputs—that is, the messages and LED states—and comparing them to the expected output. In order to comprehend the correct execution for each state, the processes can be executed step-wise (comparable to debugging a software implementation). Traces of the execution of the business processes are logged for later analysis.

The testbed allowed us to test different processes that could otherwise not have been tested by typical means of simulation. The ability to test a transaction from the start to the finish by holding a NFC tag in front of the reader of the testbed device helped us to come up with a robust solution. Errors found in an early stage could

already be ad-

[1] OCPP v1.5 WSDL - http://www.openchargealliance.org/?q=node/9

[2] RaspberryPi Model B - https://www.raspberrypi.org/products/model-b/

dressed within the next iteration (cf. CSF_3). Throughout the development phase, the testbed was used to validate the correct execution according to the OCPP. Thus, all components comprising an EV-charging infrastructure could be executed and tested in a realistic setting.

6 Discussion

6.1 Contributions to the Practice of New CPS Development

We proposed to extend the task-technology-fit perspective towards the testing of technical feasibility in the development of new CPSs, which led to the identification of "Technical Feasibility" as an *additional* critical success factor (CSF) for such projects. Motivated by the observation of OSTROM et al. that the question how service prototyping can "materialize an integrated set of service system components" [10, p. 137] has not been sufficiently addressed by service research, this paper provides an illustrative example of creating a testbed to check a CPS's technical feasibility at an early stage of service development. The testbed mimics the CPS and allows to check the CPS's business processes for technical feasibility and correct execution. We believe that this paper contributes meaningfully to the practice of new CPS development, because it shows a way to test a CPS before implementing the "real" technical components of the CPS to the full extent. Especially in highly standardized environments, our approach may enable service, product and IT engineers to align their designs not only with the business envi- ronment but also with the enabling hardware. From a managerial perspective, taking the CSF of "Technical Feasibility Fit" carefully into account can mitigate the risk and thereby the cost of erroneous specifications that would surface late in the development process. A testbed mimics the specified system and makes it accessible through a hands-on method, which enables developers to detect such erroneous specifications earlier. In regard to the demonstrated EV-charging infrastructure testbed, announced revisions of the OCPP will be addressed. The IS at hand implements business processes that have to be executed consistently across a fleet of heterogeneous charging points (e.g., different vendors, models, and revisions). Inter-compatibility can be assured using the presented method.

6.2 Contributions to Service Research

When proposing a new method, there is a need to demonstrate its "worthiness" against the existing body of knowledge. We therefore subsequently review the extant literature to demonstrate that this research creates novel scientific knowledge if transferred to contexts other than its originating one. We consider service testing in four streams in the academic literature, viz., (a) *product-service systems*, (b) *cyber-physical systems*, (c) *service marketing* and (d) *service blueprinting* as part of *service engineering*.

Cyber-physical Systems: service testing in CPS literature is mainly interpreted from a computer science point-of-view as the problem to prove the security, privacy, reliability, or resilience at the intersection of embedded computing components and cyber infrastructures [28, 29]. Thus, researchers suggest to rep- resent CPSs using formal specifications, which facilitate the adoption of formal verification techniques [29] as testing means. Formal specifications naturally focus on a CPS's physical and computational aspects.

While the CPS literature addresses interactions between physical and com- putational components only, our work also includes human interactions with physical components during the assessment of technical feasibility. Therefore, unlike the prevalent literature on testing in CPS, our work considers the entire service system comprising of people, technology, organization and information if all the CSFs are taken into account during service development.

Product-service Systems: PSS originates from manufacturing and industrial engineering with a focus on how to develop marketable customer solutions that involve physical components. PSS endeavors are often based on conventional product design processes. The properties relevant for "testing" include product-related quality issues as well as specific economic and environmental benefits of PSS [30]. While many successful PSS implementations have been reported—such as the Electrolux case study of in-flight services [31]—testing is mentioned rarely. Testing PSS primarily focuses on the physical PSS components and provides ways to step-wisely develop service concepts during iterative product design cycles [32].

Service Marketing: the prevalent metrics for assessing service concepts in service marketing relate to financial performance, anticipated market impact [33] or anticipated customer satisfaction [34]. Related to the latter, recent publications suggest techniques that strive to make a future customer's service experiencable at an early stage of service development, such as the customer journey, touchpoint approaches, and storytelling [8]. Roleplaying, design scenarios, story- boards, desktop walk-through, and service staging extend this list [35]. Addition- ally, visualizing techniques such as flow-charting, service blueprints, and process- chain-network diagrams are frequently suggested [36]. "Prototyping", however, has been widely neglected in the service marketing literature [37]. BOWERS early assumed the root-cause in arising cost for people and equipment if one wanted "to create a whole process just for testing" [38]. In the CPS setting, testing procedures from service marketing are not capable of addressing technical feasibility within a service as proposed in this study.

Service Blueprinting: the concept service blueprinting, originally proposed by SHOSTACK, is widely used in practice to analyze and design customer interaction in service systems [39]. The service blueprint depicts the division and visibility of a service system's work, structured by the actors (customers and providers) and stages (front-stage and back-stage) [40]. Blueprinting has been reported to be beneficial in new service development, management and control of existing service processes, and customer preferences monitoring (see [41]). It explicitly shows the physical evidence that is seen by the customer during various stages of service delivery. While testing back-stage activities by means of a service blueprint comes closest to what our method intends to achieve, blueprinting focuses on the non-IT components of the service system. This is why we also failed to find guidance to our problem of testing a CPS's technical feasibility in this stream of literature.

6.3 Limitations

In this paper, we proposed an enhancement to the CSFs for BPM in the context of developing new services that are enacted using CPSs. Currently, the scope of the testbed method is very limited as it is only applicable if the system behavior can be anticipated or if it is prescribed by a standard. In absence of established and validated standards, a sound testbed specification that is equivalent to the original technology is almost impossible to achieve. Therefore, future research must explicate for which scenarios the approach is suitable as our findings are derived from a very specific use case that basically dictates the technology to be used and thus cannot be generalized. Furthermore, the scope of this research is on technical feasibility. Future research should also focus on an economic evaluation of the testbed method, i.e., introduce key performance indicators that give evidence on cost savings, improved quality of business processes, and gains for the latter phases in CPS development.

7 Conclusion

High development costs make it difficult to create a full CPS prototype only for testing. However, testing is important to ensure the operational feasibility of the CPS design at an early development stage. Our work applied and extended task-technology fit in order to develop a method for creating CPS testbeds. A testbed allows to validate the correct execution of business processes while also providing evidence on the interaction of customers with the physical components. Our major contribution lays in the identification of the technological feasibility as an additional critical success factor in new CPS development, which, if considered carefully, can help to mitigate the risks of premature failure and may save costs of the CPS development and testing. We incorporated technical feasibility into a method for creating testbeds that is applicable when the behavior of the system is prescribed in standards. This method proved to be useful for our purposes in developing a new CPS for EV charging.

References

1. Tuli, K.R., Kohli, A.K., Bharadwaj, S.G.: Rethinking Customer Solutions: From Product Bundles to Relational Processes. Journal of Marketing 71(3), 1–17 (2007)
2. Spohrer, J., Kwan, S.K., Fisk, R.P.: Marketing: A Service Science and Arts Perspective. Edward Elgar Publishing (2014)
3. Vargo, S.L., Lusch, R.F.: Service-Dominant Logic: Continuing the Evolution. Journal of the Academy of Marketing Science 36(1), 1–10 (2008)

4. Johnson, S.P., Menor, L.J., Roth, A.V., Chase, R.B.: A Critical Evaluation of the New Service Development Process: Integrating Service Innovation and Service Design. Sage (2000)
5. Zeithaml, V.A., Bitner, M.J., Gremler, D.D.: Services Marketing: Integrating Customer Focus Across the Firm. McGraw-Hill, New York, NY (2012)
6. Kelly, D., Storey, C.: New Service Development: Initiation Strategies. International Journal of Service Industry Management 11(1), 45–62 (2000)
7. Cheng, C.C.: Market-Creating Service Innovation: Verification and its Associations with New Service Development and Customer Involvement. Journal of Services Marketing 26(6), 444–457 (2012)
8. Zomerdijk, L.G., Voss, C.A.: NSD Processes and Practices in Experiential Services. Journal of Product Innovation Management 28(1), 63–80 (2011)
9. Meiren, T., Burger, T.: Testing of Service Concepts. The Service Industries Journal 30 (4), 621–632 (2010)
10. Ostrom, A.L., Parasuraman, A., Bowen, D.E., Patricio, L., Voss, C.A.: Service Research Priorities in a Rapidly Changing Context. J Service Research 18(2), 127–159 (2015)
11. Geisberger, E., Cengarle, M.V., Keil, P., Niehaus, J., Thiel, C., Thönnißen Fries, H.J.: Cyber-Physical Systems: Driving Force for Innovation in Mobility, Health, Energy and Production. Springer-Verlag (2011)
12. Trkman, P.: The Critical Success Factors of Business Process Management. International Journal of Information Management 30(2), 125–134 (2010)
13. Goodhue, D.L., Thompson, R.L., Goodhue, B.D.L.: Task-Technology Fit and Individual Performance. MIS Quarterly 19(2), 213–236 (2014)
14. Sridhar, S., Hahn, A., Govindarasu, M.: Cyber-physical system security for the electric power grid. Proceedings of the IEEE 100(1), 210–224 (2012)
15. Bradley, J.M., Atkins, E.M.: Optimization and control of cyber-physical vehicle systems. Sensors 15(9), 23020–23049 (2015)
16. Mitchell, R., Chen, I.R.: Behavior Rule Specification-based Intrusion Detection for Safety Critical Medical Cyber Physical Systems. IEEE Transactions on Dependable and Secure Computing 12(1), 1–1 (2014)
17. Knight, J.C.: Safety critical systems: challenges and directions. In: Proc ICSE 2002. pp. 547–550 (2002)
18. National Academy of Science and Engineering: Cyber-Physical Systems - Driving Force for Innovations in Mobility, Health, Energy and Production. Tech. rep., acatech, Berlin, Germany (2011)
19. Hong, K.K., Kim, Y.G.: The critical success factors for ERP implementation: an organizational fit perspective. Information & Management 40(1), 25–40 (2002)
20. Tassey, G.: The Economic Impacts of Inadequate Infrastructure for Software Testing. NIST, RTI Project 7007(011), 309 (2002)
21. Lee, E.A.: Cyber Physical Systems: Design Challenges. In: Proc ISORC '08. pp. 363–369. Orlanda, FL (2008)

22. Karim, J., Somers, T., Bhattacherjee, A.: The Impact of ERP Implementation on Business Process Outcomes: A Factor-Based Study. J. Manage. Inf. Syst. 24(1), 101–134 (Jul 2007)
23. Edgar, T., Manz, D., Carroll, T.: Towards an experimental testbed facility for cyber-physical security research. Proc. of the CSIIRW '11 (2011)
24. Dijkman, R.M., Dumas, M., Ouyang, C.: Semantics and analysis of business process models in BPMN. Inf and Soft Tech 50(12), 1281–1294 (2008)
25. van der Aalst, W.M.P.: Business Process Management as the "Killer App" for Petri Nets. Software & Systems Modeling pp. 1–7 (2014)
26. van der Aalst, W.M.P.: The Application of Petri Nets to Workflow Management. Journal of Circuits, Systems and Computers 08(01), 21–66 (1998)
27. Schmutzler, J., Andersen, C.A., Wietfeld, C.: Evaluation of OCPP and IEC 61850 for Smart Charging Electric Vehicles. EV Symposium 27 (2013)
28. Wan, K., Man, K.L., Hughes, D.: Specification, analyzing challenges and approaches for cyber-physical systems (CPS). Engineering Letters 18(3) (2010)
29. Derler, P., Lee, E.a., Sangiovanni Vincentelli, A.: Modeling cyber-physical systems. Proceedings of the IEEE 100(1), 13–28 (2012)
30. Hu, H.A., Chen, S.H., Hsu, C.W., Wang, C., Wu, C.L.: Development of sustainability evaluation model for implementing product service systems. J. Environmental Science and Technology 9(2), 343–354 (2012)
31. Beuren, F.H., Gomes Ferreira, M.G., Cauchick Miguel, P.A.: Product-service systems: a literature review on integrated products and services. Journal of Cleaner Production 47, 222–231 (2013)
32. Exner, K., Sternitzke, A., Kind, S., Beckmann-Dobrev, B.: Hybrid Prototyping. In: Christoph Gengnagel, Emilia Nagy, R.S. (ed.) Rethink! Prototyping: Transdisciplinary Concepts of Prototyping, pp. 89–110. Springer (2015)
33. Johne, A., Storey, C.: New service development: a review of the literature and annotated bibliography. European Journal of Marketing 32(3/4), 184– 251 (1998)
34. Edvardsson, B., Olsson, J.: Key Concepts for New Service Development. The Service Industries Journal 16(2), 140–164 (1996)
35. Blomkvist, J.: Ways of Seeing Service: Surrogates for a Design Material. Nordes pp. 1–4 (2015)
36. Sampson, S.E.: Visualizing Service Operations. J Service Research 15(2), 182–198 (2012)
37. Blomkvist, J., Holmlid, S.: Service Prototyping According to Service Design Practitioners. Innovation pp. 1–11 (2010)
38. Bowers, M.R.: Developing new services: improving the process makes it better. Journal of Services marketing 3(1), 15–20 (1989)
39. Shostack, G.L.: How to Design a Service. E J Marketing 16(1), 49–63 (1982)
40. Becker, J., Beverungen, D., Knackstedt, R., Matzner, M., Müller, O., Pöppelbuß, J.: Bridging the Gap Between Manufacturing and Service Through IT-Based Boundary Objects. IEEE Trans Eng Man 60(3), 468–482 (2013)
41. Kostopoulos, G., Gounaris, S., Boukis, A.: Service blueprinting effectiveness: drivers of success. Managing Service Quality: an International Journal 22(6), 580–591 (2012)

How Machines are Serviced – Design of a Virtual Reality-based Training System for Technical Customer Services

Dirk Metzger[1], Christina Niemöller[1], Benjamin Wingert[2], Tobias Schultze[2], Matthias Bues[2] and Oliver Thomas[1]

[1] Universität Osnabrück, Informationsmanagement und Wirtschaftsinformatik, Osnabrück, Germany
{dirk.metzger,christina.niemoeller,oliver.thomas}@uni-osnabrueck.de
[2] Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Stuttgart, Germany
{benjamin.wingert,tobias.schultze,matthias.bues}@iao.fraunhofer.de

Abstract. Training service provider is a crucial factor for high-quality service delivery. Due to the rise of new devices, reviving Virtual Reality (VR) offer great opportunities to overcome current training challenges. As various new interaction and visualization systems push into market, guidance on how to design VR-based training systems is necessary. The presented use case is based on technicians in technical customer services (TCS) who tackle increasing complexity of machines. We fill the research gap of design knowledge by (1) analyzing the domain in a multi-method approach to elicit meta-requirements, (2) proposing design principles, and (3) instantiating them in a prototype. The interaction of the user with the training system was identified as key aspect to foster learning. We follow a design science research approach (DSR) combing the build-phase with agile evaluation cycles obtaining focus groups and demonstration with a prototype.

Keywords: Virtual Reality, Learning, Design Science Research, Technical Customer Services

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Metzger, D.; Niemöller, C.; Wingert, B.; Schultze, T.; Bues, M.; Thomas, O. (2017): How Machines are Serviced - Design of a Virtual Reality-based Training System for Technical Customer Services, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 604-618

1 Introduction

Due to the high range of tasks [1] combined with the increasing complexity of machines [2], service technicians processes of technical customer services (TCS) are complex entities. Thus, training technicians is a crucial factor for high-quality service delivery. It is common to train the relevant skills directly on the job which is time consuming and expensive as the service is completed slower or error costs for mistakes arise. Another option is the use of training facilities but they involve travel expenses, running costs for the facility and time for traveling.

Existing literature proposes the use of Virtual Reality (VR) as it provides a virtualized simulation that suits training purposes well. In addition, the immersion or presence effect of VR enhances the transfer of learned processes to the real world [3]. As VR technology is known for several years some prototypes are presented in literature (e.g. [4]) but the broad usage was, due to their costs, rather limited. With new technology and, especially, new interaction systems the usage in a larger context gets possible, which is why we revive the discussion and propose to use VR while building new training systems. Based on the example of TCS, this allows us to train the service provider (technicians) and generate a better understanding of machines with lower costs than training on the job or training facilities would provide.

The paper makes use of the design science research paradigm and follows its four step approach: (1) for analysis we used a multi-method approach to elicit requirements. (2) Based on them we derived design principles and instantiated the system. (3) We evaluated the artefact multiple times through focus groups during the development and a demonstration (formative), and (4) diffused our insights, which is done with this contribution. The core of this work is the blueprint of a VR-based training system that is based on the elicited requirements and design principles. So, further similar systems can be built analogously to our instantiation. Gregor and Hevner [5] argue that the instantiation itself contributes to the knowledge base.

The paper is structured as follows: First, we introduce the related work of VR in the next section. In section 3, we introduce our research approach. Next, we present the artefact design comprising the meta-requirements, design principles and the instantiation. We conclude by discussing novelty, practical relevance, theoretical contributions, and limitation as well as giving an outlook for future work.

2 Related Work

Virtual reality and so called head mounted displays (HMD) have a long history [6–8]. Up to the last 2-3 years most research made technical possibilities of HMD subject to discussion; only few business-related questions and use cases were discussed [9]. This was due to the fact of high cost of the first HMD systems.

On the basis of new technological developments (e.g. such as the announcement of Oculus Rift) new opportunities arise. New devices are being developed to enhance usability for the user [10]. The three principles of VR *immersion*, *interaction*, and *user* *involvement* with the environment, offers a very high potential in education by making learning more motivating and engaging [10].

Most of the research relates to medical training systems (e.g [10–13]), but also maintenance and service have been discussed (e.g. [14–19]). An example is the work of Rahimian at al. [21] who reports the use of VR for the professional training of architecture, engineering or construction specialists. The authors focus on visualization technologies to foster innovation in engineering [20]. However, the investigation of user interactions with the devices have rarely been discussed as it was not possible with the technologies so far. A close to reality training scenario could be realized only by combining the VR visualization with interaction and tracking of hands to model work steps. This is where our work starts.

Freina and Ott (2015) investigated in their state of the art of VR education several aspects for the main motivation of VR usage: it is mainly used for scenarios that "cannot be accessed physically". This limit may be due to time problems, physical inaccessibility, limits because of a dangerous situation or ethic problems [10]. In line with the authors, we started our research because of the need of maintain objects on machines that are physical inaccessible or just accessible for one technician and not the trainer. An example is the training scenario in the tank of an agricultural machine which is described in the instantiation section.

Guidance for the design of VR systems is suggested by several authors. For instance, Wann [21] took up design principles from a psychological perspective. Further, Chaturvedi et al. [22] propose core properties for building virtual worlds. Kohler et al. [23] suggest key principles such as usability, sociability, pragmatic and collaborative to build virtual worlds. Finally, Sutcliffe and Gault [24] investigated criteria for successful virtual reality applications that are suitable to be basis for design principles. All of them discussed the design in context to older virtual reality devices. Thus, we founded our design principles on them but expanded them to new aspects of new hardware and enhanced interaction components.

3 Research Approach

We follow a design science research (DSR) approach [25–27] as it is generally accepted for service systems engineering (SSE) [28]. Based upon their study about the state of the art in SSE, Böhmann et al. [28] argue, that the complex socio-technical context of service systems restricts the opportunities for meaningful laboratory-style research. Hence, they propose that research needs to be embedded within a service system in a real-world scenario and call for the design of novel service systems. In line with the authors, our approach continuously involves experts from TCS as well as observations of real-world process scenarios. We investigated the four phases analysis, design, evaluation and diffusion. By contributing our work we spread our insights (diffusion); the other three phases (analysis, design, evaluation) are shown in Figure 1.

Once the relevant business problem was defined, according to Hevner [29], attributes of the pursued future system have to be investigated and defined. These attributes are usually referred to as *meta-requirements* [30] because they reflect generic requirements that should be met by the future system (cf. section 4.1). The meta-requirements were elicited from the analysis of the real-world scenario (workshops, process analysis, expert interviews, benefit analysis).

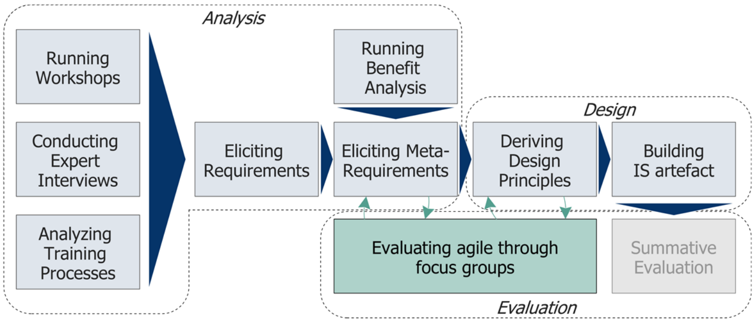


Figure 1. Research Approach

Although VR has a long scientific history, practitioners only have few experiences and just a brief idea about its usage in their specific scenario. This is why we chose to conduct a multi-method approach (workshops, expert interviews, process analysis, benefit analysis), on the one hand, to explorative analyze the domain and combine the different points of view and, on the other hand, calibrate and validate our work [31]. Next, an information system needs to be designed that meets the identified meta-requirements. Therefore, we proposed *design principles* (DPs) that describe how the new system should be built in order to fulfill the identified meta-requirements (cf. section 4.2). Finally, the IT artefact was instantiated (cf. section 4.3).

Since the evaluation of design artefacts and design theories is a central and critical part of DSR [27, 32], we combined the build-phase with several evaluation phases. How to go about choosing and designing an appropriate evaluation approach is a very significant but under-addressed issue in extant DSR literature [33]. Hence, Venable et al. [33] propose a framework for developing an appropriate evaluation strategy. Following their argumentation, our evaluation strategy is human-risk & effectiveness-oriented, due to the novelty of VR in practice. As a result, we have to evaluate our artefact early in a naturalistic setting, conducting formative and summative evaluations [33]. For implementing the evaluation strategy, we made use of agile software development [34] to rapidly develop and evaluate the software continuously (formative evaluation with focus group). So, we were able to validate the applicability to real world problems. The final summative evaluation is not part of this paper (thus, indicated in grey in Figure 1).

3.1 Requirement Collection and Analysis

User Workshops. The collection of requirements started with user workshops (first workshops). The goal was twofold. First, we wanted to get everyone's ideas on how a VR-based training system might look like (brainstorming session). Second, we wanted to collect an initial set of requirements we could start with (collection session). In order to get a broader understanding and minimize bias, we prepared the same workshop twice in two companies of different size and field. Both are part of the machinery and plant engineering sector with a major focus on TCS. The reason of splitting the workshops to different companies with different sizes was to get a company-, sector- and size-independent view. Overall, we had 13 participants from the group of technicians (user), service instructors and the management board.

For the collection of the requirements, which is the relevant part for this paper, we concluded the brainstorming session with an introduction about VR technology, its features, limitations and expectations. Afterwards, we answered everyone's questions before we asked them to write down requirements they have on a VR-based training system on paper cards. Those were put on a poster and consolidated (in case that requirements are mentioned multiple times). When everyone finished the collecting, we discussed the idea connected to every requirement. After everyone was clear, we asked them to rate every requirement (benefit analysis) on a three-tier scale ranging from 1 (very important) over 2 (important) to 3 (not that important). To minimize social influence and further specify the requirements, we asked every participant in the following interviews, as they were done with the same participants, individually what the most important requirements are.

Interviews with Practitioners. We used a semi-structured interview with eight questions separated into a general part and a requirement-specific part. In the general part, we asked questions about demographical information, their experience with VR before we started the workshop, and the general assessment of the technology. The questions were chosen to get an idea about the current situation and the attitude of the interviewee towards VR. Within the requirement-specific part, we asked the interviewee to name the two most important requirements from the workshop and about potential showstoppers that prevent people from using a VR-based training system. To sum up, the interviewee was asked to describe how training is conducted in their company, what challenges they face with the current approach and, finally, whether they think that VR might be beneficial. Regarding the documentation and the analysis procedure of the interview transcripts, the recording technique consisted of a digital record and afterwards a full text transcription and a content analysis [35].

Process Analysis. Besides the use of workshops (including interviews and benefit analysis) the current situation in training was of major importance. Therefore, we attended in multiple trainings in both companies and analyzed how technicians are taught and what role TCS processes play. Making use of shadowing [31] we participated with two IS researchers, tried to minimize our influence and documented the training itself. Afterwards, the transcripts were analyzed regarding the setting and the TCS processes to derive requirements or validate the known requirements.

Meta-Requirements. In order to generate a manageable list out of all mentioned requirements, to start implementing with, we first clustered them using open coding [35]. This was done to ensure that different aspect implied in the requirements are included. Among every clustered group, we chose the requirements that were rated highest in the benefit analysis. For the software implementation three groups were excluded as they target organizational aspects that are not relevant for the software or aspects that are not in focus of a training system. The process of generating meta-requirements was done by a researcher team of three information system researchers. Afterwards, the meta-requirements were discussed with the focus group (as described in 3.3 and appendix A). The remaining list form the (evaluated) meta-requirements that are discussed in the following.

3.2 Implementation

Starting with the elicited meta-requirements, the aforementioned team of information systems researchers derived design principles that support the design of the system. Those were based on a literature study, including papers that explicitly state design principles for Virtual Reality. Based on them the system is likely to fulfill the initial requirements. The derivation itself was executed in iterations to refine and revise the principles over and over again. When the team was convinced the meta-requirements were discussed with the focus group (as described in 3.3 and appendix A) and, thereby, evaluated to cover the most important aspects. Afterwards, the system was implemented in cooperation between the information system researchers and visual technology experts. All features were implemented in an agile approach with integration of the focus group multiple times (evaluating multiple versions of the system). The result of the implementation is described in section 4.3 and gives insights into the VR-based training system.

3.3 Evaluation

Following a human risk & effectiveness evaluation strategy, we conducted several formative evaluation cycles and a final summative evaluation [33]. For the evaluation of the meta-requirements (second workshop) and the design principles (third workshop), in each case we used discussions in form of focus groups [36, 37]. The focus group meetings were conducted in 2015 (For details about the participants see Appendix A). Different types of participants and at least two of every type were invited; on the one hand, to get different point of views, on the other hand, to have people with homogenous background to get a free-flowing conversation [36]:

1. Representing the TCS and, hence, user and customer perspective from practice, three attendees from a medium-sized service provider for air-conditioning technology and three participants from a large agricultural technology manufacturer with own TCS attended.
2. For gaining insights from a technological perspective, two IT practitioners and two visual technology researcher with expertise in implementing VR participated.
3. To bridge the technological and service view, three IS researcher specialized in service science were invited and took up the role as leader of the open discussion.
4. For the design of the content and targeted communication of information, two researchers with specialty in education and media psychology were invited.

So, with the focus group meetings, we were able to evaluate the functional design and usefulness of the VR-based training system in the business context of TCS.

4 Artefact Design

Within this section, we present the design process of our IS artefact. To specify the real world problem, we start with the meta-requirements that were generated. Based on them design principles are presented that deal as foundation of the implementation of our VR-based training system. Additionally, we give details about how our system was designed and how the design principles influenced it.

4.1 Meta-Requirements

Through the multi method-approach, we generated an overall of 69 requirements structured into seven cluster. The clustered groups were named *generating*, *content*, *interaction*, *usage*, *organization*, *output* and *general*. The groups *output*, *organization* and *usage* were excluded as they do not have a training focus but rather an organizational focus. The group *output* contains requirements concerning additional output variants of the system that might be needed in organizations but not for training. The second group *organization* targets organizational aspects that are not relevant for implementation. Finally, the third group *usage* aims at scenarios where the system might be usable besides training. With the remaining four groups 14 meta-requirements are selected that were ranked highest. They are described in the following:

- **MR1: Include CAD-data.** Usually for every machine the manufacturing company owns construction data (named Computer-Aided Design). This three-dimensional model of the machine is designed by engineers for construction purpose, before the machine itself is built. As those CAD-models (mostly) are an accurate representation of the machine, they should be the source VR builds the representation on.
- **MR2: Recording of guided trainings.** The VR-based training system should be based on a guided training that leads the trainee through the process. The trainer should be able to record or prepare trainings on his own.
- **MR3: Highlighting of parts.** As complex machines consist of several parts, the selection and highlighting of parts is beneficial to understand what parts exist and how they are installed.
- **MR4: Assembly of parts.** The system needs to simulate the assembly of parts of the machine to enable the trainee to understand which parts belong to the machine and to see parts lying underneath others.
- **MR5: Disassembly of parts.** Analogously, the disassembly of parts must be included as well to enable maintenance or repair processes.
- **MR6: Movement of parts.** While (dis-)assembling, the trainee should be able to move, rotate and inspect the parts to see and understand the details.
- **MR7: Validated interaction.** To ensure that trainees are learning the correct order on how to (dis-)assemble parts of the machine, the system needs to validate whether the interaction with the parts is possible.
- **MR8: Meta information on parts.** When inspecting or moving the parts, the system should include meta information about the part such as the name or part number.

- **MR9: Variants of machines.** As machines often exist in different variants (e.g. for different markets), the system should include information about different parts and in which variants they are used.
- **MR10: Machine parameters.** The training system should include settings and machine parameters of parts or the machine itself, as they might be crucial for the maintenance and repair.
- **MR11: Tool information.** For certain parts tools are needed (e.g. for fixing or configuring). The training system should include information about tools needed.
- **MR12: Intuitiveness.** For the training the usability plays a major role. The trainee needs fast and easy access to the system without a long training phase for the interaction with the training system itself.
- **MR13: Motivating.** On top of fast and easy interaction, the system needs to be pleasurable to use and motivating. This improves learning performance and trainee's intention to use.
- **MR14: Realistic.** Finally, the system needs to be as realistic as possible to encourage transfer of knowledge acquired in the system to a real world scenario.

Summary. Overall, we elicited 14 meta-requirements based on workshops, expert interviews and benefit analysis. In particular, the interaction with the service object (to simulate a real service scenario) is of major importance (MR3 – MR7). This is why, we explicitly focused on interaction design within our design principles in the next section.

4.2 Design Principles

Starting with the meta-requirements, we derived design principles that support the design of the training system. They are described in the following:

- **DP1: Use existing construction data.** As described, construction data for machine exist in the companies. Thus, the first design principle is to build the system on top of the construction data and integrate various data formats to allow applicability in different companies (as they might use different software for CAD). In addition to the fulfilment of MR1, another minor requirement for easy maintenance of the system benefits when the system is built on top of existing systems and software (compatible to user's domain as proposed by [24] and sustain user-created content (ES3 by [22])).
- **DP2: Integrate process data.** With the need for guided training, the reuse of already existing training processes is suggested and contributes to the principle sustain user-created content (ES3 by [22]). So, the system should be built on top of process models as main representation as they are well understood and explored in literature and practice. This would fulfill MR2 and MR7 (as the interaction can be validated by the process).
- **DP3: Integrate additional parts information.** As there are some requirements concerning information on parts, the system needs to include textual or medial information. So, when the trainee is facing or moving the particular part of the machine it is possible the learn more about it. Thus, the trainee can learn what exactly

the particular part is about. The information ranges from meta information through variants of the machine the part is used in, and parameters of the part to tools that might be needed for the particular part. This fulfills MR8-MR11. The design principle was not present in literature about VR design principles [21–24]. However, it was explicitly asked by the technicians in the workshops and by the education and media psychology researchers in the focus group evaluation.

- **DP4: Carefully design parts interaction.** As different interaction with parts of the machine is needed, this becomes a main aspect of the system. For the training system the interaction is the key feature that enables trainees to understand and learn. Thus, special attention about the way of interacting with the system is crucial (in line with principle 1 by [21] and the usability principle by [23]). The design principle contributes to MR3-MR6.
- **DP5: Reduce interaction possibilities.** Finally, recent interaction systems in VR offer a wide range of pattern (such as controller-based, hand-based, voice-based etc.). However, from the requirements and gained experience the last design principle is, to limit the actually used interaction with the system to one primary approach. This ensures that users are not overwhelmed by different concepts, patterns and interaction approaches. Consequently, this simplifies the familiarization of users with the system. From our point of view, we encourage the usage of hand-recognition-based interaction as it appears to be the most natural way to interact with the system (in line with [24] and [23]). Thus, it fulfills MR12-MR14 as it improves the system's usability, motivation to use and, when hand-recognition is used, the realism of the system.

Summary. Overall, we derived five key design principles. Each of them contributes to at least one meta-requirement. Table 1 summarizes all five design principles and their relation to the meta-requirements and literature.

Table 1. Design Principles

DP	Description	MR	Literature
1	Use existing construction data	1	[24][22]
2	Integrate process data	2,7	[22]
3	Integrate additional parts information	8-11	
4	Carefully design parts interaction	3-6	[21][23]
5	Reduce interaction possibilities	12-14	[24][23]

4.3 Instantiation

The system components (see Figure 2, numbered from 1-3) of the VR-based training system are primarily based on Oculus Rift DK2 (see 1) as a display and tracking device. The tracking captures orientation and position of the head relative to the Oculus Camera. Thus, the interaction in VR are precise and the visualization immersive. For further user interaction, a Leap Motion controller (see 3) is mounted in front of the Oculus head mounted display (HMD). The Leap Motion is a high frame per second hand tracking camera device and gives the opportunity to display hand and motion fluid into the visual experience of the user. To enhance the hand interaction area, a new 30-degree angle mount (see 2) was built to shift the hand area down and support a more natural hand input environment (fulfilling DP4). For the rendering and visual input, a high end gaming computer is needed to realize the necessary power for the experience.

For the visualization pipeline, an own linear learning authoring tool was developed that combined the CAD data (fulfilling DP1) with processes (fulfilling DP2) and results in an immersive VR learning experience. The authoring tool allows using CAD data without ever preprocessing it for the VR environment. On the one hand, the data is not visually optimized to display in VR, but on the other hand, a fast and cost efficient way to create new learning scenarios without touching the CAD data is a big advantage in the creation of lessons.

Figure 2. Hardware component setup

Figure 3. User hand avatar

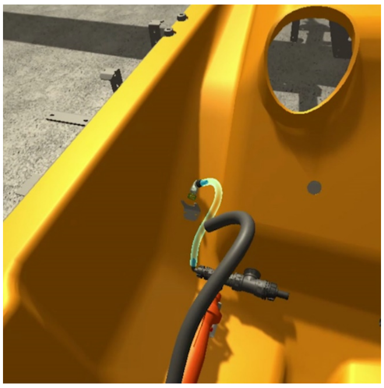


Figure 4. Placing components

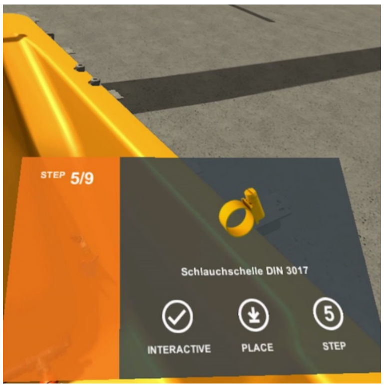


Figure 5. User manual and menu

In the VR-based training system, the whole interface and interaction techniques are mapped to hand gestures and hand movement (see Figure 3). Complex finger tasks are simplified to give the user a better control and robust recognition. The learning tasks are more oriented towards the knowledge transfer than learning certain hand movements (see Figure 4) (fulfilling DP4 and DP5). Additionally, we built a user manual and menu interface (see Figure 5) that comes up as soon as the system detects the inside of the users hand. So, the interaction is twofold: with the out-/upper side of the hand the user can interact with the simulation and on the in-/inner side he or she gets an interface with additional information (fulfilling DP3).

The system is designed for standing users with approximately 2x2m interaction space. Due to the need of a high end gaming computer and the amount of components it is best used in an (mostly) immobile setup.

5 Conclusion, Discussion and Outlook

Conclusion. The effective use of Virtual Reality (VR) offers great opportunities to overcome current challenges in the domain of TCS. Due to the complexity of service systems engineering [28], guidance on how to design service support systems is necessary. To overcome this complexity and fill the research gap of design knowledge on VR-based training systems, we followed a DSR approach within this paper through, first, exploring the domain and eliciting meta-requirements (step 1), and second, deriving design principles continuously working in an interdisciplinary team of practitioners and researchers (step 2). The key feature was the design of the interaction with the service object, as it was the major requirement (to simulate a real service scenario).

Discussion. Before we started the project on VR, our main expectation was that for users the most important requirements would concern the acceptance of the VR device itself. We thought of motion sickness, refusal of wearing the device on the head or problems with spectacle wearers. Surprisingly, the most important requirements we found was about the interaction with the system. So, for our users the VR itself seems to be acceptable straight away as major companies (e.g. Facebook, Sony, HTC) are pushing into market. So, users get more and more in contact with the technology and used to the idea of virtual worlds. We experienced broad interest through all companies and individuals we talked to. Most of them mentioned deployment scenarios in their own processes that might be useful (e.g. constructors (teaching order of assembly), engineers (visualizing unbuilt machines), farmers (introduction and commissioning to tractors)). We further found scenarios that are not trainable without VR, due to danger-related (e.g. the repair of running wheel gears; to build an understanding about what happens when certain gears are damaged) or size-related reasons (e.g. the repair of parts inside of a tank of a spraying machine that is just the size of one person).

Not only because of the broad interest, one main aspect we found was that integration of potential users in development is very crucial. New technologies are connected to a learning process on both sides, developer and user, which is why they have to talk about possibilities and see what works for them and what does not. This is how we came up with integrating a gesture interaction component to offer the hand-based interaction as it was the most natural option for users.

Finally, the VR-based training is suitable at least to teach scenarios that are not trainable otherwise. Whether it might help in other scenarios as well and improve the training is a research question that arose during the project for further investigation.

Novelty and Practical Relevance. We address a real-world problem consisting of need for TCS support through training systems that teach complex machines. At the same time, since recent VR is still an emerging technology, little knowledge about the design of training systems exist. Both from the point of practice and from theory, a transfer of the proposed design knowledge to other user groups or even customers offer new subjects of research. Thus, we formulated the design principles to be as generic and applicable as possible to other user groups and sectors. With our instantiation we demonstrated how to build a VR-based training system using recent hardware (Oculus Rift DK2 and Leap Motion). Hence, cost-efficient training with commercial VR technology is possible, which is of major relevance for small and medium sized companies.

Theoretical Contribution. Regarding the theoretical contribution, this research work contributes to the methodological knowledge base of IS Design and service systems engineering, and builds upon existing methods of DSR and the design of service systems. In DSR, a theoretical contribution is usually regarded to be in form of prescribing how a specific solution can be designed in order to solve a relevant real-world problem; often presented in form of design principles [38, 39] that guide the implementation of specific instantiations. Gregor and Hevner [5] argue that the instantiation itself contributes to the knowledge base as the demonstration of a novel artefact can be a research contribution that embodies design yet to be articulated, formalized, and fully understood. Prescriptive knowledge can be generated through (1) inventing new solutions for new problems, (2) improving and thereby developing new solutions for existing problems, or (3) adopting known solutions to solve new problems [5]. We position our work as a new solution, a VR-based training system, to solve an existing problem consisting of complex machines that needs to be serviced by TCS. We build our work on the knowledge base of Virtual Reality Design Principles. However, for the field of technical customer service the integration of additional service related information and the design of the interaction with the service object is crucial. Thus, we enhanced the known principles with our study. Therefore, we explored the problem domain and formulated meta-requirements that represent the conditions that should be met by a VR trainings solution. Additionally, we contribute to the IS research knowledge base by instantiating the evaluation strategy proposed by Venable et al. [33] in combination with agile software development as enhancement of the classic DSR approach. Hence, with our work, developed in a transdisciplinary team obtaining IS research, service science, education and media psychology as well as practitioners from service providers, manufacturers and IT companies, we meet a research gap and the claim for evidence-based design research [28].

Limitations and Outlook. We discussed our work with experts from two different sectors in order to transfer the design principles to other sectors. However, researchers are welcome to evaluate the applicability and potentially needed adoption separately as the VR-based training system has a wide area of possible applications. Based on the results of the benefit analysis, we focused on an excerpt of all collected requirements and the evaluation of the VR-based training system with technicians. Hence, (1) the transfer of additional requirements have to be investigated further. (2) We have not conducted a summative evaluation regarding the economic benefit and the training success yet. Thus, the next step of our research is the evaluation of our instantiation in form of a field test [40] in the TCS of the agricultural technology company and the service provider for air-conditioning technology. Thereby, an experiment that allows conclusions about the training effect is needed. To sum up, our approach can be considered as VR with new hardware that might lead research to specify new business models or training options.

6 Acknowledgement

This paper is part of the project *Glassroom* and is funded by the German Federal Ministry of Education and Research (BMBF, grant number 01PD14014A).

References

1. Baines, T., Lightfoot, H., Smart, P., Fletcher, S.: Servitization of manufacture - Exploring the deployment and skills of people critical to the delivery of advanced services. Journal of Manufacturing Technology Management. 24, 637–646 (2013)
2. Däuble, G., Özcan, D., Niemöller, C., Fellmann, M., Nüttgens, M.: Design of User-Oriented Mobile Service Support Systems - Analyzing the Eligibility of a Use Case Catalog to Guide System Development. In: Thomas, O. and Teuteberg, F. (eds.) 12. Internationale Tagung Wirtschaftsinformatik (WI 2015). 149–163. Osnabrück (2015)
3. Regenbrecht, H.T., Schubert, T.W., Friedmann, F.: Measuring the sense of presence and its relations to fear of heights in virtual environments. International Journal of Human-Computer Interaction. 10, 233–249 (1998)
4. Seymour, N.E., Gallagher, A.G., Roman, S.A., O'Brien, M.K., Bansal, V.K., Andersen, D.K., Satava, R.M.: Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Annals of surgery. 236, 458–463 (2002)
5. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. Management Information Systems Quarterly. 37, 337–355 (2013)
6. Rheingold, H.: Virtual Reality: Exploring the Brave New Technologies. Simon & Schuster Adult Publishing Group (1991)
7. Caudell, T.P., Mizell, D.W.: Augmented reality: an application of heads-up display technology to manual manufacturing processes. In: Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences, 1992. 659–669 vol.2 (1992)
8. Sutherland, I.E.: A head-mounted three dimensional display. In: Proceedings of the December 9-11, fall joint computer conference, part I. 757–764. ACM, NY, USA (1968)
9. Teuteberg, F.: Mobile Augmented Reality aus betriebswirtschaftlicher Sicht. HMD - Praxis der Wirtschaftsinformatik, Mobile Anwendungen. 42, 86–94 (2005)
10. Freina, L., Ott, M.: A literature review on immersive virtual reality in education: state of the art and perspectives. In: The International Scientific Conference eLearning and Software for Education. p. 133 (2015)
11. Faria, J. de, Teixeira, M.: Virtual and stereoscopic anatomy: when virtual reality meets medical education. Journal of neurosurgery (2016)

12. Aïm, F., Lonjon, G., Hannouche, D., Nizard, R.: Effectiveness of Virtual Reality Training in Orthopaedic Surgery. The Journal of Arthroscopy (2016)
13. Huang, H., Liaw, S., Lai, C.: Exploring learner acceptance of the use of virtual reality in medical education: A case study of desktop and projection-based display systems. Interactive Learning Environments (2016)
14. Webel, S., Bockholt, U., Engelke, T., Gavish, N., Olbrich, M., Preusche, C.: An augmented reality training platform for assembly and maintenance skills. Robotics and Autonomous Systems. 61, 398–403 (2013)
15. Westerfield, G., Mitrovic, A., Billinghurst, M.: Intelligent Augmented Reality Training for Assembly Tasks. Artificial Intelligence in Education. 542–551 (2013)
16. Seth, A., Vance, J.M., Oliver, J.H.: Virtual reality for assembly methods prototyping: a review. Virtual Reality. 15, 5–20 (2010)
17. Gavish, N., Gutierrez, T.: Design guidelines for the development of virtual reality and augmented reality training systems for maintenance and assembly tasks. BIO Web of Conferences. 29, 1–4 (2011)
18. Yuviler-Gavish, N., Krupenia, S., Gopher, D.: Task Analysis for Developing Maintenance and Assembly VR Training Simulators. Ergonomics in Design: The Quarterly of Human Factors Applications. 21, 12–19 (2013)
19. Woll, R., Damerau, T., Wrasse, K., Stark, R.: Augmented reality in a serious game for manual assembly processes. 37–39 (2011)
20. Pour Rahimian, F., Arciszewski, T., Goulding, J.: Successful education for AEC professionals: case study of applying immersive game-like virtual reality interfaces. Visualization in Engineering. 2, 4 (2014)
21. Wann, J., Mon-Williams, M.: What does virtual reality NEED?: human factors issues in the design of three-dimensional computer environments. International Journal of Human-Computer Studies. 44, 829–847 (1996)
22. Chaturvedi, A.R., Dolk, D.R., Drnevich, P.L.: Design Principles for Virtual Worlds. 35, 673–684 (2011)
23. Kohler, T., Fueller, J., Matzler, K., Stieger, D., Füller, J., Matzler, K., Stieger, D.: Co-creation in virtual worlds: The design of the user experience. MIS Quarterly. 35, 773–788 (2011)
24. Sutcliffe, A., Gault, B.: Heuristic evaluation of virtual reality applications. Interacting with Computers. 16, 831–849 (2004)
25. Peffers, K., Tuunanen, T., Rothenberger, M. a., Chatterjee, S.: A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems. 24, 45–77 (2007)
26. Österle, H., Becker, J., Frank, U., Hess, T., Karagiannis, D., Krcmar, H., Loos, P., Mertens, P., Oberweis, A., Sinz, E.J.: Memorandum on design-oriented information systems research. European Journal of Information Systems. 20, 7–10 (2011)
27. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Quarterly. 28, 75–105 (2004)
28. Böhmann, T., Leimeister, J.M., Möslein, K.: Service Systems Engineering. Business & Information Systems Engineering. 6, 73–79 (2014)
29. Hevner, A.R.: A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems. 19, 87–92 (2007)
30. Walls, J.G., Widmeyer, G.R., El Sawy, O. a: Assessing information system design theory in perspective: How useful was our 1992 initial rendition. Journal of Information Technology Theory & Application. 6, 43–58 (2004)

31. Myers, M.: Qualitative Research in Business & Management. Sage Publications Ltd., London (2009)
32. March, S.T., Smith, G.F.: Design and natural science research on information technology. Decision Support Systems. 15, 251–266 (1995)
33. Venable, J., Pries-heje, J., Baskerville, R.: FEDS: a Framework for Evaluation in Design Science Research. European Journal of Information Systems. 1–13 (2014)
34. Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R.C., Mellor, S., Schwaber, K., Sutherland, J., Thomas, D.: Manifesto for Agile Software Development, http://agilemanifesto.org/, (2001)
35. Mayring, P.: Qualitative Inhaltsanalyse Grundlagen und Techniken. Beltz, Weinheim (2010)
36. Morgan, D.L.: Focus Groups as Qualitative Research. SAGE Publications (1996)
37. Krueger, R.A., Casey, M.A.: Focus Groups: A Practical Guide for Applied Research. SAGE Publications (2014)
38. Kuechler, W., Vaishnavi, V.: A Framework for Theory Development in Design Science Research: Multiple Perspectives Science Research. Journal of the Association for Information Systems. 13, 395–423 (2012)
39. Sein, M.K., Henfridsson, O., Rossi, M., Lindgren, R.: Action Design Research. MIS Quarterly. 35, 37–56 (2011)
40. Sonnenberg, C., vom Brocke, J.: Evaluations in the Science of the Artificial - Reconsidering the Build-Evaluate Pattern in Design Science Research. DESRIST 2012, LNCS 7286. 381–397 (2012)

Table 2. Appendix A: Workshop participants

#	*Position*	*Sector*	*Experience in position*	*Size of the company*
1	Research assistant	University, Information Systems	3 years	~ 1.700
2	Research assistant	University, Information Systems	3 years	~ 1.700
3	Professor	University, Information Systems	7 years	~ 1.700
4	Managing director	Service provider for air-conditioning	> 20 years	~ 240
5	Assistant to the board of directors	Service provider for air-conditioning	2 years	~ 240
6	Assistant to the board of directors	Service provider for air-conditioning	4 years	~ 240
7	Research assistant	Application-oriented research, Visual Technology (VR expert)	1 year	~ 560
8	Researcher, Head of Visual Technology	Application-oriented research, Visual Technology (VR expert)	> 10 years	~ 560
9	Service trainer	Agricultural technology	> 10 years	~ 1.800
10	Team Assistant After Sales	Agricultural technology	1 year	~ 1.800
11	Head of After Sales	Agricultural technology	> 20 years	~ 1.800
12	Researcher	University, Education and Media Psychology	> 20 years	~ 2.800
13	Professor	University, Education and Media Psychology	> 20 years	~ 2.800
14	IT expert	IT company, Learning Technologies	5 years	~ 220
15	IT expert	IT company, Learning Technologies	1 year	~ 220

Multi-Channel Choice in Retail Banking Services: Exploring the Role of Service Characteristics

Dennis Hummel[1], Silvia Schacht[1], Alexander Mädche[1]

[1] Karlsruhe Institute of Technology (KIT), Institute of Information Systems and Marketing (IISM) and Karlsruhe Service Research Institute (KSRI), Karlsruhe, Germany
{dennis.hummel}@partner.kit.edu, {silvia.schacht, alexander.maedche}@kit.edu

Abstract. Companies are moving towards omni-channel management offering all products and services on all channels. Yet, some of these investments might be obsolete as certain products are associated with certain channels. At this point, service companies are still left behind as past research focused on product categories and it remains unclear if the results are transferrable to services. Our study addresses this gap by analyzing the influence of service characteristics on channel choice. We tested our research model by surveying 2,000 banking customers in Germany on their past channel choices for five financial services. The results show that complex services with a high value are rather purchased in a branch than the online channel. Thereby, demographics and behavior-related constructs are important control variables. The results improve the understanding of channel choice behavior in a multi-channel context for services and provide guidance for practitioners to right-channel IT investments.

Keywords: Multi-Channel Choice, Omni-Channel Management, Service Characteristics, Retail Banking Services, Multinomial Logistic Regression

1 Introduction

Recent research suggests that companies are moving from a multi-channel to an omni-channel management which seems to be the next generation of channel management [1]. Omni-channel management comes with several advantages: It allows customers to seamlessly shift between channels and use them simultaneously, and it improves the user experience [1]. To move from multi- to omni-channel capabilities, serious IT investments are required. For example, Commerzbank, the second largest bank in Germany, is investing more than 200 million Euro into their multi-channel banking platform to make services accessible on all channels [2]. Thereby, it is debatable whether all products and services have to be offered on each channel (omni-channel), or whether certain products and services are more suitable to particular channels than others (multi-channel). For instance, Gupta et al. [3] find that search goods are more likely to be purchased online, whereas experience goods are more likely to be purchased offline.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Hummel, D.; Schacht, S.; Maedche, A. (2017): Multi-Channel Choice in Retail Banking Services: Exploring the Role of Service Characteristics, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 619-633

In general, the role of product categories on user's multi-channel choice is well established in retailing (e.g. [3, 4]). Yet, it is unclear if the results are transferrable to services. While products are tangible, exchangeable, tradable and preservable, services typically fulfill the opposite of characteristics: Intangible, heterogeneous, inseparable and perishable [5]. Yet, there is no commonly shared definition of services [5], and bundles of products and services further complicate the differentiation. Nevertheless, most multi-channel studies have focused on clearly defined products (such as books) and services (e.g. [6–8]). Gupta et al. [3] examine the role of product categories for three products and one service. They find varying significances of price search intentions on channel switching behavior for the different product categories. Similarly, Gensler et al. [7] find different results for two product categories in retail banking that could not be explained theoretically. In particular for financial services, they suggest to examine the role of service categories in greater detail in future research. Also other researchers suggest to research the role of product categories on channel choice [9]. Thereby, it is important to note that product *categories* are "all the products offering the same general functionality" [10] whereas product *characteristics* are features of a product like color, weight, price, etc. Services can also be categorized along characteristics and be grouped into service categories. Thus, we aim to answer the following research question:

Do service characteristics influence user's channel choice across all stages of the buying process in the context of a multi-channel environment in financial services?

To address this question, we build on the multi-channel literature from peer-reviewed papers which report channel behavior in a multi-channel context. Based on the literature, we developed a research model and surveyed banking customers in Germany on their past channel choices for five financial services. Due to the mixed results regarding the effects of demographics on individuals' channel choice, we additionally controlled for several demographic characteristics and experiences.

The contribution of this work is manifold as we consider service characteristics of five different financial services across four different stages of the buying process. Moreover, we study different channels using a strong empirical foundation of 2,000 banking customers. In addition, we focus on the users' demographics and the often neglected importance of service characteristics on channel choice. Our research also comes with implications for practitioners: Following an omni-channel strategy of providing all products and services on all channels without a clear objective can result in waste of scarce resources. Instead, managers should carefully balance which products and services are offered on which channel. Our study is to be distinguished from other studies as it addresses channel choices of purchases rather than tasks (e.g. [11]) and it is conducted in a multi- and not a single-channel environment.

The remainder of this paper is organized as follows: Chapter 2 recaps the related work. In chapter 3, we describe the constructs and hypotheses to answer our research question. Chapter 4 outlines the data collection and the applied research methodology. In chapter 5, we summarize the results of our research and discuss them in more detail in chapter 6. Finally, we conclude our research in chapter 7.

2 Related work

Some studies consider product categories as a determinant of channel choice (e.g. [3, 7, 12, 13]). For an overview, we provide a summary of 10 studies in Table 2. Several insights can be derived from the analysis: Almost all studies note that product categories and characteristics have an influence on user's channel choice (e.g. [7, 13]), customer segments (e.g. [14, 6]), and customer value (e.g. [4]). Only partially, product categories seem to have no impact as, for instance, Maity and Dass [12] could not find an influence of the product type on information search. Further, the related work reveals that researchers developed various characteristics to classify categories, namely complexity, perceived risk, purchase frequency, search/experience goods, etc. In addition, it becomes apparent from the analysis that many different channels and product categories have been examined, but hardly any services. Focusing on services, Table 1 displays studies from the context of financial services. The comparison shows that service characteristics are only one covariate of channel choice and that there is a broader picture behind channel choices in a multi-channel environment. Beyond the service characteristics, the stage of the buying process, the channel characteristics, and the demographics influence user's channel choice. For instance, Gensler et al. [7] suggest that the stage of the buying process has a moderating impact on channel choice, and many other researchers also included several stages in their studies (e.g. [3, 15, 16]). This paper addresses all aspects except for the already well-studied channel characteristics such as price, risk or convenience (e.g. [17]).

Table 1. Related work on multi-channel choices in financial services

Source	*Channels (among others)*	*Several stages*	*Channel Charact.*	*Demo-graphics*	*Service Charact.*	*Services*
[18]	Branch, Internet, phone	-	X	X	X	1; 5-8
[19]	Branch, Internet	X	X	X	-	-
[20]	Branch, Internet, ATM	-	X	-	-	-
[16]	Branch, Internet	X	X	X	X	6
[21]	Branch, Internet, phone	X	-	X	-	1-4; 6; 8
[7]	Branch, Internet, phone	X	X	X	-	1; 2

1: Checking account; 2: Securities account; 3: Credit; 4: Savings account; 5: Retirement provision; 6: Mortgage; 7: Insurance; 8: Investments

In financial services, only Black et al. [18] and Frambach et al. [16] considered the service characteristics. The former study is an explorative, qualitative work that yielded valuable insights into channel choices in services but that, to the best of our knowledge, has not yet been validated by practice. The latter study is an empirical analysis but only includes one service (mortgage) and thus it does not allow for comparisons between different services. Other studies of channel choice in financial services do not account for service characteristics (e.g. [19–21]). Overall, we can derive from the related work that product and service characteristics influence user's channel choice.

Table 2. Subset of related work on product categories in a multi-channel context

Study	*Channel*	*Categories*	*Results*
[18]	1, 2, 3	Financial products	- Type of financial product a key influencer for channel selection - Product description along complexity and perceived risk
[15]	1, 2, 4, 5, 7	10 broad categories	- Different channels used for searching different product categories - Assessment of shopping attributes differs for product categories and channels - Categorized along purchase frequency, durability and degree of entertainment
[22]	1, 2, 4	Fashion	- Product-channel associations (e.g. store for personal items; Internet for functional repeat purchases)
[13]	1, 2, 4	11 categories	- Hypothesized about products-channel-associations; investigated channel migration along product type - Results show that product categories generally influence customer's channel choice
[3]	1, 2	Flight tickets, books, wine, stereo systems	- Drivers and inhibitors of channel switching differ across product categories - Categorization along search and experience goods - Search goods more likely to be purchased online, whereas experience goods more likely offline
[6]	2, 3, 4	Household and personal products	- Multichannel behavior across product categories; categories influence of consumer segments - Classification along complexity, purchase frequency and tangibility - Covariates of multichannel shopping (i.e. channel choice) differ across product categories
[7]	1, 2, 3, 6	Checking account, brokerage account	- Results indicate that effects of channel choice attributes could differ between product categories - Assumption that product categories might explain differences and suggest to undertake further research
[23]	1, 2, 4	Furniture, appliances, music, books, etc.	- Generally, significant relation between product categories and cross-channel free-riding behavior - Free-riding behavior differs across respective product categories
[12]	1, 2, 5	Airline tickets, food delivery	- Product type moderates effect of media richness on perceived fit, satisfaction, and channel choice - No support for the moderating effect of product type on information search
[14]	1, 2, 3	Mobile solutions (telecomm.)	- Replication of [6]; perceived product complexity determines affiliation to certain customer segments - Covariates of multichannel shopping differ across categories
[24]	1, 2	Apparel, consumer electronics	- Product category affects usage of on- vs. offline channels in each stage - Shoppers more involved with a category use Internet (maybe due to increased purchase frequency)

1: Store/branch; 2: Internet; 3: Telephone/Call Center; 4: Catalog; 5: Mobile; 6: ATM; 7: Other

3 Development of constructs and hypotheses

After scanning existing literature on the effects of the most common service characteristics, we realized that there is quite some research in the field of products, but only few results on services. Thus, based on the related work, we formulate a set of hypotheses that can be tested for services rather than products. We focus our hypotheses on the two main banking channels that are also apparent from the related work section: the branch and the online channel.

3.1 Service complexity

Service complexity is a well-known construct to characterize a service. The construct is adapted from product complexity and is defined as "the extent to which the consumer perceives a service to be difficult to understand or use" [22:112]. Complexity of products is addressed by several researchers (e.g. [14, 6, 16, 18]). Black et al. [18] recognize the importance of complexity of financial products on channel choice. Even advanced Internet users in their focus groups had strong reservations for purchasing complex services online, and they preferred to go to the store [18]. Frambach et al. [16] also study service complexity and find that people prefer the offline channel for complex services. They argue that customers need help in their decision making and thus opt for the branch. Finally, Konus et al. [6] find that the Internet is too complex for shopping, whereas this finding could not be supported in a replication study [14].

Hypothesis 1 (H1): For services with a high complexity, customers prefer the branch, while for services with a low complexity, they prefer the online channel.

3.2 Service value

Service value describes the monetary value of a service purchased. Service value is closely associated with perceived risk, because when the involved sums are high, then there is a greater probability of making a bad decision [26]. Burke [15] supports this finding by stating that customers are more likely to visit the store when looking for expensive and infrequently purchased goods. Service value also impacts the number of channels used. Heitz-Spahn [23] argues that customers engage in more searching when the product or service is expensive and thus automatically use more channels. Sullivan and Thomas [13] find that channel choice varies by the total amount of money spent.

Hypothesis 2 (H2): For services with a high value, customers prefer the branch, while for services with a low value, they prefer the online channel.

3.3 Service purchase frequency

Purchase frequency describes the "total number of purchase occasions" [12:16]. Purchase frequency has been studied in the context of product categories by various researchers (e.g. [13, 15, 23, 27]). Customers tend to choose the store when purchasing goods infrequently, because they need advice from knowledgeable sales people [15]. Sullivan and Thomas [13] find a relationship between channel choice and purchase frequency as customers which use the Internet and the catalog buy more frequently than customers which choose the store.

Hypothesis 3 (H3): For services with a high purchase frequency, customers prefer the online channel, while for services with a low purchase frequency, they prefer the branch.

3.4 Service usage frequency

To our knowledge, service usage frequency has not been studied in the context of multi-channel choices. Gensler et al. [7] suggest studying usage frequency to explain differences between financial products. For our research, we interpret service usage frequency as the frequency a customer is using or interacting with the service after its purchase. A comparable way of looking at usage frequency is the product return behavior of customers. Several researchers have shown that the more often customers return products, the more likely they use a digital channel and vice versa (e.g. [28]).

Hypothesis 4 (H4): For services with a high usage frequency, customers prefer the online channel, while for services with a low usage frequency, they prefer the branch.

3.5 Stage-channel associations

Finally, we study stage-channel associations. Stage-channel associations “exist when consumers associate a certain stage of the buying process with a particular channel” [5:989]. Thereby, stages represent the various phases of the buying process. Channels, like the branch or the Internet, are the contact points that customers employ for information, contracting, and use. For example, the Internet is often associated with the search stage, while the store is associated with the purchase stage [8]. Other researchers find similar stage-channel associations (e.g. [15, 22, 27]), but it is unclear if they also exist in the context of services.

Hypothesis 5 (H5): The buying process stage has a moderating impact on users’ channel choice as the information stage is associated with the online channel while the purchase stage is associated with the branch.

3.6 Demographics, Internet usage and Internet experience

The influences of demographics on channel choice in a multi-channel environment are questionable. Some researchers argue that there is no relationship between demographics and channel choice (e.g. [6, 21]). Older studies, however, find that age [18, 29, 30], income [18], education [29], and gender [30] influence individuals’ channel choice. Thereby, young, male customers with high income and education favor the Internet. Therefore, we control our data for demographical differences.

Beyond demographics, we also control for the influence of behavior-related constructs (e.g. Internet experience, Internet usage, online banking usage, etc.) on channel choice. It is not surprising that Internet experience favors the channel choice of the Internet channel (e.g. [16]). Yet, we have found no other studies that consider the effect of online banking usage and mobile banking usage on channel choice in a multi-channel environment. In addition, no other study considered the effect of low experiences on online banking in a multi-channel context.

The hypotheses introduced above are integrated in a research model (see Figure 1).

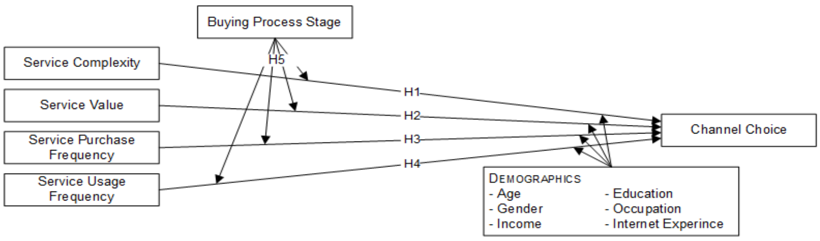


Figure 1. Research model

4 Research methodology

4.1 Data collection

To test our hypotheses, we build on a dataset that was collected and analyzed by Graupner et al. [31]. The authors employed a market research firm that conducted a survey among 2,012 German retail banking customers from May to June 2015. The survey was hosted online and assessed the actual past channel choice for different financial services. Intentionally, customers without Internet access were excluded from this study as they are not able to use the online channel of this study. The financial services under observation were credit, checking account, savings account, securities account, and retirement provision and each service is represented at least 300 times in our survey. Although theory of planned behavior states that behavioral intentions are antecedents for actual behavior [32], asking for actual behavior is advantageous over intentions to use a channel as intended behavior may not result in actual behavior. The survey covered seven different channels that are the most common ones in German retail banking: branch, online, mobile, telephone, postal, family and friends, and other. For the definitions of the respective channels see Hoehle et al. [33]. Family and friends are not a banking channel per se but a way to inform oneself or to receive counseling. We have chosen the sector of financial services for our study as it offers a wide range of services to compare and as multiple channels are already available since decades [33].

Table 3 displays the demographics. The sample of our data collection comprises more men than women with an average age of 42 years. 29% of the respondents report an average income between 1,500€ and 2,499€ per month and 70% are employed. The sample is representative for the German Internet population of 2015.

Table 3. Demographics

Attribute		*Total*	*%*	*Attribute*		*Total*	*%*
Gender	Male	1,066	53%	**Income**	< 1,500€	521	26%
	Female	946	47%		1,500-2,499€	590	29%
Age	18-29	410	20%		2,500-3,499€	502	25%
	30-39	550	27%		> 3,500€	399	20%
	40-49	405	20%	**Education**	In training	190	9%
	50-59	370	18%		Working	1,405	70%
	>60	277	14%		Not working	417	21%

4.2 Measurement of dependent variable

The dependent variable is the channel choice of the customer for the information, counseling, contracting, and usage stage for the respective financial service. During the information stage, customers search for information (e.g. by scanning the website) while during the counseling stage they are actively seeking advice from a person, for instance in the branch or in the online channel (e.g. video chat with an advisor). For the contracting and usage stage see Gensler et al. [7]. In the survey, the respondents specified the chosen channel for one service in each stage of the purchasing process. Channel choice is a categorical variable as there is no hierarchy between the channels.

Table 4 illustrates the channel choices across the different stages of the buying process. It displays that the branch and the online channel are by far the most often chosen channels in our data set as they cover 70% of all choices. Thus, we will focus our results and the subsequent discussion on these two channels.

Table 4. Channel Choices across different stages of the buying process

Channel	*Information*	*Counseling*	*Contracting*	*Usage*[*]	*Total choices*
Branch	755 (38%)	1,038 (52%)	1,183 (59%)	437 (20%)	3,413 (35%)
Online	939 (47%)	546 (27%)	633 (31%)	719 (33%)	2,837 (35%)
Mobile	34 (2%)	22 (1%)	16 (1%)	40 (2%)	112 (1%)
Phone	20 (1%)	54 (3%)	34 (2%)	127 (6%)	235 (3%)
Postal	25 (1%)	15 (1%)	59 (3%)	127 (6%)	226 (3%)
Friends	157 (8%)	99 (5%)	n/a	n/a	256 (3%)
Other	82 (4%)	53 (3%)	87 (4%)	17 (1%)	239 (3%)
None	n/a	185 (9%)	n/a	684 (32%)	869 (11%)

[*] Multiple answers possible

4.3 Measurement of independent variables

The independent variables are the service characteristics of service complexity, value, usage frequency and purchase frequency. The assessment of the independent variables was not conducted by the survey participants themselves. Instead, the values were gathered by asking ten employees of different seniority and from different departments of a large German commercial bank for their expert assessment. This approach was used to reach an objective and transferrable evaluation of the respective financial services. An evaluation by the survey participants might have caused a social desirability bias towards rating their own choice (i.e. financial service) as particularly difficult and complex to manage. The bank employees were given a short questionnaire where they had to specify the service characteristics for each financial service on a 5-Point-Likert scale ranging from 5 ("very high") to 1 ("very low"). The inter-rater reliability of their judgements was low and explanations could solve all disagreements. For the analysis, the average of the different responses was used for each characteristic rather than the most frequent entry of the assessment to allow for more fine-grained differences in the evaluation of the respective dimensions.

4.4 Data analysis

Based on the related work, we assume that customers are influenced in their channel choice during the different stages of the buying process based on service complexity, value, usage frequency and purchase frequency. In addition, the stage is hypothesized to have an effect on the dependent variable. Thus, the channel choice is expressed by the following formula (1):

$$CC_{c,s,t} = \alpha_{c,s,t} + \beta_{c,s,t} \times Complex + \gamma_{c,s,t} \times Value + \delta_{c,s,t} \times Use_{frequ} + \eta_{c,s,t} \times Purch_{frequ} + \chi_s \times Stage \quad (1)$$

whereas,	$CC_{c,s,t}$	Channel choice of customer c for service s in stage t
	$\alpha_{c,s,t}$	Constant for channel c for service s in stage t
	$\beta_{c,s,t}$	Factor for service complexity for channel c for service s in stage t
	$\gamma_{c,s,t}$	Factor for service value for channel c for service s in stage t
	$\delta_{c,s,t}$	Factor for service purchase frequency for channel c for service s in stage t
	$\eta_{c,s,t}$	Factor for service usage frequency for channel c for service s in stage t
	χ_s	Factor for stage of the purchasing process for service s

As the dependent variable is a nominal variable, we used a multinomial logistic regression (MLR) method to calculate the research model. Thereby, we employed the statistical software R and especially the nnet package that can be used for Feed-Forward Neural Networks and Multinomial Log-Linear Models. The MLR is an extension of the binary logistic regression by estimating separate binary models for each category. In the end, only N-1 binary logistic regression models are displayed as one category serves as a reference and each binary model estimates the effect in comparison to this reference category (in our case the online channel).

5 Results

First, Table 5 shows the dimensions of service characteristics. For example, the securities account and the retirement provision are rated much more complex than the checking or savings account. Moreover, checking and savings account are purchased and used more often than the retirement provision.

Second, we examined the distribution of the channel choices. Figure 2 shows the relative amount of the channel choices across the four stages of the buying process. The graphs are similar in their structure. Most customers inform themselves online with a range of 33% (retirement provision) to 54% (checking account) of the customers. The branch is used equally or less often at this stage. A major shift occurs in the counseling stage when more customers visit the branch. For all financial services, the branch is predominant (between 48% and 56%). Only for the checking account, the difference between the online channel and the branch is small (9%).

Table 5. Dimensions of service characteristics

Service	*Complexity*	*Value*	*Purchase frequency*	*Usage frequency*
Credit	3.1	3.9	2.6	2.6
Checking account	1.5	2.0	5.0	3.6
Savings account	1.7	2.6	2.5	2.5
Securities account	3.9	3.7	3.0	2.7
Retirement provision	4.1	3.6	2.0	2.0

1=Very low; 2=Low; 3=Medium; 4=High; 5= Very high

Also in the contracting stage, most customers prefer the branch over other channels. The difference is particularly visible when studying the retirement provision as only 18% of the customers used the online channel to purchase the credit compared with 63% using the branch. Finally, most customers again use the online channel for the after-sales stage. Yet, here the data scatters from 37% (retirement provision) to 66% (checking account).

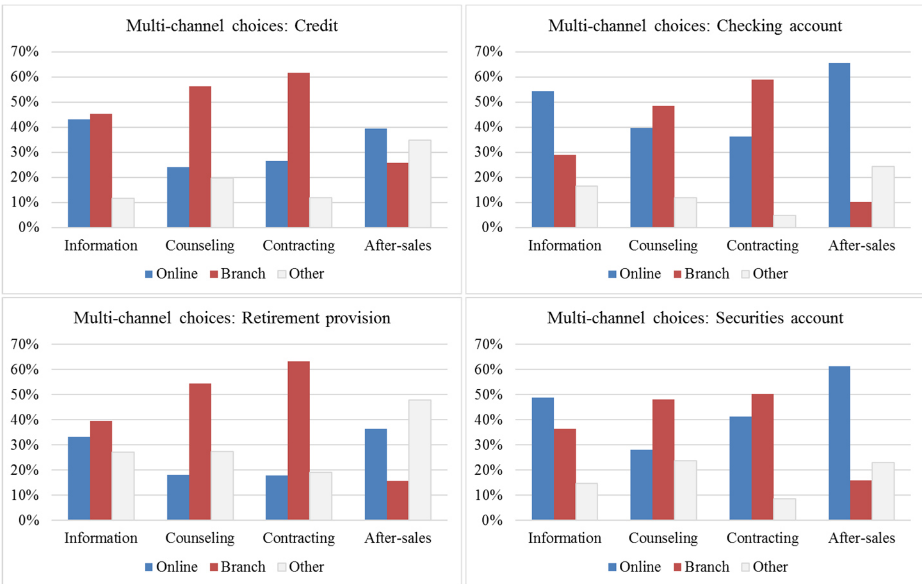


Figure 2. Channel choices of different financial services across stages of the buying process

In a next step, we analyzed whether the above mentioned differences, especially in the contracting and the after-sales stage, are significant or within ranges that can still be explained by coincidence. Table 6 displays the results of the MLR. Although the complete results are displayed, we only discuss the branch and the online channel.

The results show that the channel choice is significantly different in the branch and the online channel in the contracting stage (all characteristics significant) and the after-sales stage (only complexity and value significant). For instance, value of services yield in a different channel choice for the branch ($p<0.01$) compared with the online channel in the contracting stage. Further, complexity leads to a different channel choice between branch ($p<0.04$) and the online channel in the after-sales stage. The direction of the effects are consistent across all stages. Complexity and usage frequency always carry a negative sign, while the effect for value and purchase frequency is positive. We find more significant differences for other channels. Yet, the results for them are not reliable due to small case numbers (see Table 4).

Moreover, we estimate the effect of the stage of the purchasing process on the channel choice. Again, we used a MLR for the analysis and the results are shown in Table 6. There exists a relationship between the stage of the buying process and the chosen channel as the channel choice is significantly different for the branch and the online channel. Together with Table 4, we can state the online channel is associated with the information and after-sales stage, while the branch is associated with the counseling and the contracting stage.

Table 6. Results for stage-channel associations (H5b)

Channel	*Stage*	*Channel*	*Stage*
Branch	**-0.21 (0.00)**	FF[1]/ ATM	**-0.42 (0.00)**
Mobile	-0.09 (0.34)	Other	**-0.61 (0.00)**
Telephone	**0.21 (0.01)**	None	**0.27 (0.00)**
Postal	**0.76 (0.00)**		

Effects of buying process stage on channel choice. p-values in brackets. Values in bold are significant at the 5% level
1: Family and Friends

When controlling for demographics and behavior-related constructs, all demographics are significant at least in one stage, e.g. age (counseling and contracting stage), gender (all stages) or education (all stages except for counseling stage). This implies, for example, that women are using the branch significantly more often than men across all stages of the buying process. In addition, education is influential as customers with higher education are more likely to choose the online channel and they are less likely to draw upon counseling. On the other hand, channel choices seem to be almost independent of income and occupation. A similar picture arises from behavior-related constructs, where online banking usage and low experiences with the online banking are significant in all stages, whereas Internet usage is only significant in the information and contracting stage. It is evident that customers having low experiences with online banking are more inclined to avoid the online channel. The opposite holds true for online banking usage.

6 Discussion

We started from the question whether service characteristics influence channel choices. Our results show that the channel choice between the branch and the online channel is significantly different for the constructs of complexity, value, and usage frequency in the contracting and the after-sales stage. They are not significant in the first two stages. Thus, we can partly accept H1, H2, H3 and H4 as complexity and value influence channel choices in the last two stages while purchase frequency and usage frequency are only significant in one stage.

Table 7. Results of the research model across the different stages of the buying process

Stage	*Information*				*Counseling*			
Channel	*Complex*	*Value*	*Purch_fre*	*Use_fre*	*Complex*	*Value*	*Purch_fre*	*Use_fre*
Branch	-0.07 (0.77)	0.39 (0.31)	0.41 (0.53)	-1.13 (0.36)	-0.23 (0.41)	0.68 (0.10)	1.21 (0.08)	-2.43 (0.07)
Mobile	-0.84 (0.45)	0.92 (0.57)	1.78 (0.52)	-3.31 (0.55)	0.14 (0.81)	-0.07 (0.95)	0.21 (0.81)	-0.88 (0.50)
Telephone	1.17 (0.22)	-1.17 (0.43)	-3.47 (0.21)	6.49 (0.22)	0.29 (0.65)	0.44 (0.67)	-1.35 (0.50)	3.67 (0.36)
Postal	-0.73 (0.18)	1.08 (0.25)	**3.07 (0.00)**	**-6.66 (0.00)**	-1.20 (0.11)	1.39 (0.22)	**2.06 (0.02)**	**-4.67 (0.00)**
FF[1]	0.58 (0.25)	-0.80 (0.32)	1.01 (0.38)	-2.20 (0.33)	1.14 (0.06)	-1.34 (0.18)	0.29 (0.84)	-0.68 (0.80)
Other	-0.31 (0.62)	1.60 (0.11)	2.68 (0.13)	-6.20 (0.05)	-0.85 (0.32)	2.69 (0.06)	**5.49 (0.03)**	**-11.22 (0.01)**
None	n/a	n/a	n/a	n/a	-0.16 (0.70)	0.55 (0.39)	0.56 (0.61)	-0.77 (0.72)
Stage	*Contracting*				*After-sales*			
Channel	*Complex*	*Value*	*Purch_fre*	*Use_fre*	*Complex*	*Value*	*Purch_fre*	*Use_fre*
Branch	**-0.89 (0.00)**	**1.46 (0.00)**	**2.85 (0.00)**	**-5.95 (0.00)**	**-0.79 (0.03)**	**1.36 (0.01)**	0.46 (0.62)	-1.88 (0.30)
Mobile	0.51 (0.46)	-0.16 (0.90)	-0.24 (0.82)	0.52 (0.71)	-1.32 (0.30)	2.14 (0.26)	3.52 (0.29)	-5.71 (0.39)
Telephone	-0.64 (0.44)	1.31 (0.30)	1.86 (0.39)	-4.67 (0.26)	**-1.51 (0.02)**	**3.16 (0.00)**	3.42 (0.06)	-6.54 (0.61)
Postal	**-1.95 (0.01)**	**2.89 (0.01)**	**4.56 (0.02)**	**-9.56 (0.02)**	**-1.14 (0.02)**	**2.96 (0.00)**	**4.10 (0.00)**	**-9.00 (0.00)**
ATM	n/a	n/a	n/a	n/a	-1.20 (0.39)	0.74 (0.70)	0.90 (0.79)	-2.21 (0.75)
Other	**-1.33 (0.03)**	**3.35 (0.00)**	**6.48 (0.00)**	**-12.88 (0.00)**	-0.47 (0.21)	**2.36 (0.00)**	-0.04 (0.97)	**-2.98 (0.00)**
None	n/a	n/a	n/a	n/a	**-3.99 (0.00)**	**6.90 (0.00)**	**8.84 (0.00)**	**-18.59 (0.00)**

Number of observations: 2.012; Effects of service complexity, value, purchase frequency and usage frequency on channel choice. p-values in brackets. Values in bold are significant at the 5% level

1: Family and Friends

H5 is fully supported as there is a significant effect of the stage on the channel choice. In addition, demographics and behavior-related constructs moderate the influence on channel choice, e.g. as males and customers with higher education favor the Internet channel. Thereby, our results confirm the idea that the channel choice is dependent on the purchased services and its characteristics.

Hence, our research is in line with prior studies on product categories which showed that the type of product or service influences the channel choice (e.g. [7, 13]). We agree with Black et al. [18] that banking customers rarely choose the online channel for complex services and we can confirm the findings of Burke [15] that customers prefer the store for expensive goods. Additionally, we can confirm the findings of other researchers (e.g. [15, 27, 8]) that certain stages are associated with certain channels. Finally, we find a significant influence of demographics on channel choice and thus disagree with some researchers (e.g. [6, 21]) in this regard.

7 Summary and conclusion

This work studies the influence of service characteristics on channel choice. Based on previous literature, we developed a research model including four service characteristics, and tested it across five financial services in the retail banking industry in Germany. Our results show that channel choices for the branch and the online channel are significantly different for the service characteristics during the contracting and the after-sales stage. Additionally, there is a relationship between the stage of the buying process and the channel chosen. In general, our results are in line with prior studies that examined product characteristics. Practitioners can use the results to determine which of their services to offer on which channel. We suggest digitizing less complex services with a low value, especially in the contracting stage. Further, practitioners can use stage-channel associations to tailor the channels according to customers' preferred contact points in each stage of the buying process.

However, this paper has some limitations. First, the channel choices are retrospective and only allow for a limited view on how customers purchase services in the future. Additionally, many banks did not offer all services online so that banking customers did not have a free choice across all channels. Second, the assessment of the service characteristics was performed by banking professionals (more objective but more distant from channel choices) and not the users (more subjective but more relatable to channel choices) themselves. It is debatable which assessment yields more meaningful results. Third, we neglected the channel characteristics that might have an influence on channel choices, too. Fourth, customers without Internet access were excluded as the survey was conducted online. Despite a high internet user penetration in Germany, this might create a bias towards the online channel. Finally, the paper only addresses the retail banking sector in Germany and the results might not be transferable to product-service bundles, other industries, or countries.

Thus, future research might replicate our study in other service industries (e.g. telecommunications, consulting or education) or other countries (e.g. United States or Austria). Moreover, the interesting results on demographics could be taken into account in future studies.

References

1. Verhoef, P.C., Kannan, P.K., Inman, J.J.: From Multi-Channel Retailing to Omni-Channel Retailing. Introduction to the Special Issue on Multi-Channel Retailing. J. Retail. 91, 174–181 (2015).
2. Obertreis, R.: Ein neues Konto in sieben Minuten, http://www.tagesspiegel.de/wirtschaft/commerzbank-experimentiert-ein-neues-konto-in-sieben-minuten/13746416.html.
3. Gupta, A., Su, B., Walter, Z.: An Empirical Study of Consumer Switching from Traditional to Electronic Channels : A Purchase-Decision Process Perspective. Int. J. Electron. Commer. 8, 131–161 (2004).
4. Kushwaha, T., Shankar, V.: Are Multichannel Customers Really More Valuable? The Moderating Role of Product Category Characteristics. J. Mark. 77, 67–85 (2013).
5. Parry, G., Newnes, L., Huang, X.: Goods, Products and Services. In: Service Design and Delivery. pp. 19–29 (2011).
6. Konus, U., Verhoef, P.C., Neslin, S.A.: Multichannel Shopper Segments and Their Covariates. J. Retail. 84, 398–413 (2008).
7. Gensler, S., Verhoef, P.C., Böhm, M.: Understanding consumers' multichannel choices across the different stages of the buying process. Mark. Lett. 23, 987–1003 (2012).
8. Verhoef, P.C., Neslin, S.A., Vroomen, B.: Multichannel customer management: Understanding the research-shopper phenomenon. Int. J. Res. Mark. 24, 129–148 (2007).
9. Sonderegger-Wakolbinger, L.M., Stummer, C.: An agent-based simulation of customer multi-channel choice behavior. Cent. Eur. J. Oper. Res. 23, 459–477 (2015).
10. Reibstein, D.: Product Category, http://kwhs.wharton.upenn.edu/term/product-category/.
11. Meuter, M.L., Bitner, M.J., Ostrom, A.L., Brown, S.W.: Choosing Among Alternative Service Delivery Modes: An Investigation of Customer Trial of Self-Service Technologies. J. Mark. 69, 61–83 (2005).
12. Maity, M., Dass, M.: Consumer decision-making across modern and traditional channels: E-commerce, m-commerce, in-store. Decis. Support Syst. 61, 34–46 (2014).
13. Sullivan, U., Thomas, J.S.: Customer Migration: An Empirical Investigation Across Multiple Channels. Univ. Illinois Urbana - Champaign Northwest. Univ. 41 (2004).
14. De Keyser, A., Schepers, J., Konus, U.: Multichannel customer segmentation: Does the after-sales channel matter? A replication and extension. Int. J. Res. Mark. 32, 453–456 (2015).
15. Burke, R.R.: Technology and the Customer Interface: What Consumers Want in the Physical and Virtual Store. J. Acad. Mark. Sci. 30, 411–432 (2002).
16. Frambach, R.T., Roest, H.C.A., Krishnan, T.: The Impact of Consumer Internet Experience on Channel Preference and Usage Intentions Across the Different Stages of the Buying Process. J. Interact. Mark. 22, 1–12 (2008).
17. Neslin, S.A., Grewal, D., Leghorn, R., Shankar, V., Teerling, M.L., Thomas, J.S., Verhoef, P.C.: Challenges and Opportunities in Multichannel Customer Management. J. Serv. Res. 9, 95–112 (2006).
18. Black, N.J., Lockett, A., Ennew, C., Winklhofer, H., McKechnie, S.: Modelling

consumer choice of distribution channels: an illustration from financial services. Int. J. Bank Mark. 20, 161–173 (2002).

19. Montoya-Weiss, M.M., Voss, G.B., Grewal, D.: Determinants of Online Channel Use and Overall Satisfaction with a Relational, Multichannel Service Provider. J. Acad. Mark. Sci. 31, 448–458 (2003).
20. Albesa, J.G.: Interaction channel choice in a multichannel environment, an empirical study. Int. J. Bank Mark. 25, 490–506 (2007).
21. Cortinas, M., Chocarro, R., Villanueva, M.L.: Understanding multi-channel banking customers. J. Bus. Res. 63, 1215–1221 (2010).
22. Nicholson, M., Clarke, I., Blakemore, M.: "One brand, three ways to shop": situational variables and multichannel consumer behaviour. Int. Rev. Retail. Distrib. Consum. Res. 12, 131–148 (2002).
23. Heitz-Spahn, S.: Cross-channel free-riding consumer behavior in a multichannel environment: An investigation of shopping motives, sociodemographics and product categories. J. Retail. Consum. Serv. 20, 570–578 (2013).
24. Frasquet, M., Mollá, A., Ruiz, E.: Identifying patterns in channel usage across the search, purchase and post-sales stages of shopping. Electron. Commer. Res. Appl. 14, 654–665 (2015).
25. Burnham, T.A., Frels, J.K., Mahajan, V.: Consumer Switching Costs. A Typology, Antecedents, Consequences. 31, 109–126 (2003).
26. Schoenbachler, D.D., Gordon, G.L.: Multi-channel shopping: understanding what drives channel choice. J. Consum. Mark. 19, 42–53 (2002).
27. Inman, J.J., Shankar, V., Ferraro, R.: The Roles of Channel-Category Associations and Geodemographics in Channel Patronage. J. Mark. 68, 51–71 (2004).
28. Kumar, V., Venkatesan, R.: Who are the multichannel shoppers and how do they perform?: Correlates of multichannel shopping behavior. J. Interact. Mark. 19, 44–63 (2005).
29. Strebel, J., Erdem, T., Swait, J.: Consumer Search in High Technology Markets: Exploring the Use of Traditional Information Channels. J. Consum. Psychol. 14, 96–104 (2004).
30. Venkatesan, R., Kumar, V., Ravishanker, N.: Multichannel Shopping: Causes and Consequences. J. Mark. 71, 114–132 (2007).
31. Graupner, E., Trenz, M., Maedche, A.: Understanding Digitization across Processes and their Phases – An Extension of Process Virtualization Theory. Under Rev. Eur. J. Inf. Syst. 1–34 (2015).
32. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Decis. Process. 50, 179–211 (1991).
33. Hoehle, H., Scornavacca, E., Huff, S.: Three decades of research on consumer adoption and utilization of electronic banking channels: A literature analysis. Decis. Support Syst. 54, 122–132 (2012).

Digital Nudging: Altering User Behavior in Digital Environments

Tobias Mirsch[1], Christiane Lehrer[1], Reinhard Jung[1]

[1] University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland
{tobias.mirsch,christiane.lehrer,reinhard.jung}@unisg.ch

Abstract. Individuals make increasingly more decisions on screens, such as those on websites or mobile apps. However, the nature of screens and the vast amount of information available online make individuals particularly prone to deficient decisions. Digital nudging is an approach based on insights from behavioral economics that applies user interface (UI) design elements to affect the choices of users in digital environments. UI design elements include graphic design, specific content, wording or small features. To date, little is known about the psychological mechanisms that underlie digital nudging. To address this research gap, we conducted a systematic literature review and provide a comprehensive overview of relevant psychological effects and exemplary nudges in the physical and digital sphere. These insights serve as a valuable basis for researchers and practitioners that aim to study or design information systems and interventions that assist user decision making on screens.

Keywords: Digital Nudging, Choice Architecture, Behavioral Economics, Human-Computer Interaction, User Interface Design

1 Introduction

Human decision making is imperfect. Research in psychology and behavioral economics has shown that individuals are influenced by various psychological effects during their decision making – consciously or unconsciously [1]. In fact, decisions are highly context-dependent; that is, they are influenced by the choice environment [2]. The reliance on heuristics and the influence of psychological effects such as social norms lead individuals to make predictable mistakes and often decide to their own detriment. Against this background, Thaler and Sunstein introduced the concept of libertarian paternalism as an approach to deliberately design choice environments to affect human behavior while respecting individual freedom of choice. Libertarian paternalism aims at helping individuals make better decisions in their own interest [2]. Choice environments can be designed using so-called nudges, which are relatively minor changes to decision environments. Nudges either attempt to overcome or use specific psychological effects to guide individuals towards a predefined choice option. Nudges refer to "any aspect of the choice architecture that alters individuals' behavior in a predictable way without forbidding any options or significantly changing their

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Mirsch, T.; Lehrer, C.; Jung, R. (2017): Digital Nudging: Altering User Behavior in Digital Environments, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 634-648

economic incentives” [2, p. 6]. Designing choice environments through the purposeful implementation of nudges is called nudging. A prominent example for nudging in the physical sphere is the change of cafeteria design to guide students towards a healthier diet without eliminating unhealthy foods from the menu. This is achieved by positioning healthy food options at eye level, thus making them easier to reach compared to unhealthy options [2]. In research, various disciplines, such as medicine [e.g., 3], psychology [e.g., 4], and different areas from sociology [e.g., 5] have dealt with the concept of nudging. The literature mainly discusses the application of nudging in the development of policies [e.g., 6], encouraging environmentally friendly behavior [e.g., 7], and promoting healthy lifestyles [e.g., 8]. In practice, nudging has been picked up by a number of companies and governments, which increasingly try to influence individuals’ choices [9].

The concept of nudging is increasingly gaining relevance in the digital sphere, as nowadays more and more decisions are taken on screens, such as websites or mobile apps, ranging from the choice of a travel destination to purchases of all types to the right life partner, insurance, or investment. However, in the digital environment, individuals are particularly prone to making deficient decisions. Due to the vast amount of information available on the Internet, individuals often fail to process all the relevant details to reach an optimal choice. Instead, individuals often make decisions on screens in a hasty and automated manner [10]. In this context, nudging can be an effective tool to guide users’ decision making. Compared to physical contexts, digital environments provide several advantages for nudging: the implementation of digital nudges is easier, faster and cheaper; moreover, the Internet provides specific functionalities, like user tracking, which allows personalization of nudges presented to users, making them potentially more effective [11].

While nudging has gained momentum in various fields of research as well as in practice, digital nudging has not gained much attention by information systems (IS) scholars. Against this background, we present digital nudging as a relevant and fruitful research area for IS research and for human-computer interaction (HCI) research in particular. However, prior HCI research has used behavioral economics concepts that focused mainly on a few selected heuristics and biases, such as the endowment effect, loss aversion [e.g., 12], or the status quo bias [e.g., 13]. In behavioral economics, Benartzi and Lehrer [10], and in IS, Weinmann et al. [11] extend the nudging concept to the digital context. Weinmann et al. [11] define digital nudges as user interface (UI) design elements that affect choices and propose a five-step process for developing nudges in online decision environments (see figure 1).

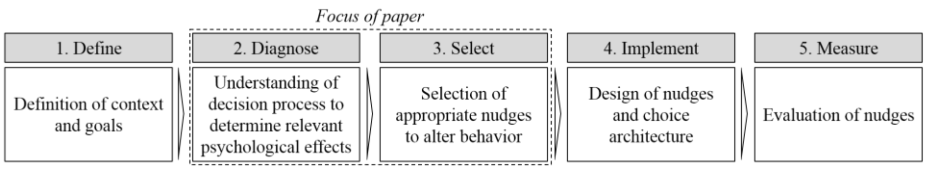


Figure 1. Nudging Development Process [11]

In the present paper, we heed the call of Weinmann et al. [11] for further research to gain a sound understanding of the mechanisms that underlie nudging. To achieve this research goal, we conducted a systematic literature review encompassing research from different disciplines. As a result, we provide an overview of relevant psychological effects that have been discussed in relation to nudging. Moreover, we present examples of digital nudges to illustrate possible approaches in practice. Thereby we address the second and third steps of the nudging development process, i.e., "Diagnose" and "Select". Based on these steps, concrete nudges and choice architecture can be developed ("Implement") and tested in lab experiments or in real world settings ("Measure"). With regard to the first step ("Define"), we do not limit our literature review, as well as the nudges presented, to a specific digital context, but provide a broad range of possible application areas.

The contribution of this paper is twofold. First, from a research perspective, we present digital nudging as a relevant and promising research area in the IS, particularly the HCI domain. In this paper, we provide an overview of the body of knowledge regarding relevant psychological effects that underlie nudges in the physical context. Thus, we illuminate the theoretical mechanisms that may also be at play in digital nudging. The psychological effects and nudges presented provide a valuable basis for behavioral researchers who aim to transfer them to the digital context and empirically examine their effects on user behavior. Moreover, our findings can guide design-oriented researchers when designing IS and interventions that assist users in making self-beneficial choices. Second, for practitioners, the concept of digital nudging provides new stimuli for UI and user experience (UX) design. A deeper understanding of the psychological effects at play in human decision making and behavior helps UI designers intentionally develop theoretically based nudges. By doing so, they can either make use of a specific psychological effect to reach a certain goal (e.g., increase sales or transaction speed) or counteract its influence. The exemplary digital nudges provide initial ideas as to how they may be implemented. Moreover, as all UI design decisions influence user behavior, UI designers can use the knowledge about the effects to verify if the current choice environment of their IT artefacts nudges users in the intended way or not.

This paper proceeds as follows. First, we present the theoretical background of behavioral economics, nudging, and HCI. Subsequently, the methodology of the literature review and the results are presented. The paper concludes with a summary, limitations, and proposals for further research.

2 Theoretical Background

2.1 Behavioral Economics and Nudging

Traditionally, economics views the human being as homo economicus, whose decision making is fundamentally rational. However, this view disregards behavioral studies of cognitive and social psychology that have empirically shown that humans do not always behave and decide rationally [2]. Behavioral economics combines psychology and economics to investigate and model human behavior with consideration for cognitive limitations and complications. Thereby, “behavioral economics increases the explanatory power of economics by providing it with more realistic psychological foundations” [14, p. 1].

According to dual process theories, dominant in the field of social psychology, individuals use different cognitive systems to assess information during the decision making process: on the one hand, there is intuitive System 1, which is fast, automatic, effortless and emotionally charged; and on the other hand, there is reason-based System 2, which is slower, effortful and deliberately controlled [15]. Most empirical studies in the field have concluded that everyday activities are mainly driven by System 1, making human decision making prone to heuristics and biases [15, 16]. Heuristics, i.e., simple rules of thumb, facilitate and accelerate the decision making process by reducing the amount of information processed. Moreover, the external environment, or choice context, is an important parameter in the decision making process [e.g., 17]. For example, different contexts may alter the assessment of trade-offs or comparisons between different options.

Nudging is a concept based on insights from behavioral economics aiming to alter environments in a way that would increase the likelihood of certain behaviors. A nudge is a simple intervention within the choice architecture to steer individuals by addressing specific psychological effects to make use of or overcome them. What differentiates nudges from other forms of intervention is that they are designed to preserve full freedom of choice [2]. Nudges are, for example, notifications that inform individuals of their calorie intake, nutrition labels on food or the automatic enrollment in a pension plan with an opt-out option [18]. Stipulating a certain diet or exercise or enrolling someone without an opt-out option would not be considered a nudge. Transferred to the digital context, digital nudging refers to the "use of user-interface design elements to guide people's choices or influence users' inputs in online decision environments" [11, p. 3]. These UI design elements include graphical design, specific content, wording or small features (e.g., product ratings) [11].

2.2 UI Design in the HCI Domain

Research in the field of HCI studies and designs interfaces facilitating the interaction between users and IT artefacts, such as websites, applications, or devices. UI design aims at maximizing the usability and UX [19]. The usability of an IT artefact refers to its ease of use and efficiency. UX can be associated with various meanings, ranging from “traditional usability to beauty, hedonic, affective or experiential aspects of technology use” [20, p. 91]. According to Hassenzahl and Tractinsky, “UX is about technology that fulfils more than just instrumental needs in a way that acknowledges its use as a subjective, situated, complex and dynamic encounter” [20, p. 95]. It can be described as a consequence of the internal state of the user, including, for example, needs, motivation, expectations, or feelings.

HCI scholars have provided various principles and guidelines for good UI design [e.g., 21, 22]. Those guidelines are based on a sound understanding of individuals’ behavior and needs and acknowledge demographic diversity as a starting point for the design process (e.g., IFIP reference model [23]). Due to the heterogeneity and changes in how humans interact with IT artefacts, UI design principles do not represent ultimate laws for design. In fact, HCI research continuously tries to advance its approaches to improve interfaces and experiences in relation to technological and user development. In doing so, HCI research often leans on insights from other fields, such as ethnography or even phenomenological philosophy [24]. We claim that behavioral economic insights and the concept of nudging are inspirations for HCI research. First, research is informed by a real-world phenomenon: the imperfection of human decision making and how relatively simple it could be addressed with digital nudges. Second, through this approach, the gap between theory and practice can be bridged by providing a first analysis. Third, it can represent an approach to discover and develop new theories as well as empirical methods or an understanding of how different approaches may complement each other. Overall, a basis for further discussion of underlying issues or support to draw conclusions from experiments with empirical results can be established. Through this approach, HCI researchers may be able to provide UI designers with insights and guidelines to increase performance or user satisfaction and lower error rates [25].

3 Literature Review on Nudging

3.1 Systematic Literature Review

To provide a comprehensive overview of the existing research on nudging, the underlying psychological effects as well as related areas, such as libertarian paternalism and behavioral economics in the digital context, we conducted a literature review in April 2016. Following the methodology proposed by vom Brocke et al. [26], we performed a search spanning multidisciplinary databases providing access to academic journals and conference proceedings. We conducted four searches by applying relevant phrases (see table 1) in the fields title, keywords, and abstract.

Table 1. Results of the literature review

Search Phrase / *Database*	*Nudging / Nudge*	*Choice Architecture*	*Libertarian Paternalism*	*Behavioral Economics AND Online*
ScienceDirect	506	1232	14	13
EbscoHost	167	652	46	4
AISeL	1	21	0	2
Unfiltered results	673	1884	60	19
Sum of relevant articles	65			

From these results, we excluded duplicates and articles not published in journals or conferences. Afterwards, we screened the articles to evaluate if they contributed to this paper. During this process, we excluded articles not topic-related, for example, articles about improving ozone modelling using observational nudging in a prognostic meteorological model or articles about the impact of nudging coefficient for the initialization on the atmospheric flow field and the photochemical ozone concentration. In a last step, we selected those articles that report on concrete nudges or/and psychological effects. For example, some articles just reported on the acceptance of nudging in society but did not elaborate on the underlying psychology or exhibit examples. After this evaluation, we considered 65 articles to be relevant for this work. Table 1 provides a detailed overview of the results.

3.2 Identified Psychological Effects and Nudges

Through the literature review, we identified a total of 20 psychological effects in the context of libertarian paternalism and nudging. Most articles described the underlying psychological effects and the associated nudges as well as a concrete application or illustrated example. However, some papers only reported psychological effects without providing examples of nudges, while others reported nudges without touching upon underlying psychological effects. In the latter case, we complemented the described nudges with the psychological effects based on gained expertise and insights. Table 2 provides an overview of the identified psychological effects based on the literature review. The frequency of appearance is higher than the number of identified papers because many papers referred to more than one psychological effect.

Table 2. Psychological effects extracted from literature

Psychological effects	*Frequency*	*Works reported on effect*
Framing	34	[4], [6-8], [18], [27-55]
Status Quo Bias	30	[3], [6-8], [18], [28], [32], [37], [42], [44], [49-65]
Social Norms	15	[5], [7], [18], [28], [37], [39], [42], [44], [64], [66-71]
Loss Aversion	13	[6], [32], [34], [35], [37], [42], [64], [66], [71-75]
Anchoring & Adjustment	7	[28], [35], [42], [50], [64], [71], [75]
Hyperbolic Discounting	7	[18], [32], [44], [64], [71], [76], [77]
Decoupling	6	[18], [32], [37-39], [77]
Priming	6	[28], [34], [64], [75], [78]
Availability Heuristic	5	[6], [44], [64], [71], [75]
Commitment	4	[6], [18], [36], [64]
Mental Accounting	4	[28], [64], [75], [79]
Optimism & Over-Confidence	4	[35], [64], [71], [77]
Attentional Collapse	3	[18], [32], [77]
Messenger Effect	3	[39], [64], [80]
Image Motivation	2	[45], [64]
Intertemporal Choice	2	[18], [71]
Representativeness & Stereotypes	2	[71], [75]
Endowment Effect	1	[75]
Spotlight Effect	1	[81]

Academic literature has mainly discussed nudging in relation to promoting healthy and environmentally friendly behavior. With regard to health, the authors discuss and empirically investigate nudges that influence food choices through framing effects such as labels, which indicate the healthiness of food [e.g., 30, 47, 48], or the positioning of healthy food options in an easily accessible way in cafeterias and/or increased visibility [e.g., 41, 53, 54]. With regard to environmentally friendly behavior, research examined nudges using social norms, such as messages that refer to the mass by stating, for example, that 70% of customers purchased at least one ecological product [5]. Furthermore, research discussed nudges based on loss aversion (e.g., subsidizing less polluting or taxing polluting travel options) [66] and anchoring and adjustment (e.g., setting reference points to evaluate eco-friendliness) [35].

The following section describes the identified psychological effects and the associated examples of nudges in more detail. Additionally, we provide examples of possible approaches for nudges in digital contexts. For this purpose, we selected well-known websites. Still, we do not claim that the examples of digital nudges are the result of a purposeful implementation by the UI designers based on the nudging concept. Nevertheless, they carry psychological effects and can be observed as nudges. These examples mainly serve to illustrate how digital nudges may appear in practice. Before providing a detailed description of every psychological effect, it must be mentioned that they partly overlap [2]. Additionally, as highlighted by Thaler and Sunstein, nudges rarely ground on only one specific psychological effect but rather on the interplay of a few different effects [2]. Furthermore, due to the length restrictions of this paper, we focused on the most frequently mentioned psychological effects (i.e., framing, status quo bias, social norms, loss aversion, anchoring & adjustment, hyperbolic discounting, decoupling, priming, and availability heuristic).

Framing. Tversky and Kahneman describe the term framing as the act of designing a decision frame in a way that the "decision-maker's conception of the acts, outcomes, and contingencies associated with a particular choice" [82, p. 453] is governed through psychological principles. By this means, shifts and outcomes of decisions are more predictable and probabilities are altered. Framing refers to a controlled presentation of a decision problem considering different framing methods regarding one decision problem. In this paper, we follow this definition but focus specifically on accentuation, orientation, and presentation of decision problems. A vivid example retrieved through the literature review shows how to reduce accidents on curvy roads by painting a series of white stripes on the streets (horizontal to the driving direction). The stripes alter the perception of speed for drivers – the driven speed was perceived as faster than it really was. Therefore, the drivers intuitively slowed down, and accidents were reduced [43]. In this example, the perception of speed was framed through a targeted accentuation and different (perceived) presentation of the environment, which altered the probability to reduce the speed. In the digital context, a practical application example can be observed on Amazon.com. On the product pages, Amazon accentuates product-related items. In doing so, the choice architecture is intervened by pulling the attention of the user to related articles. This accentuation may trigger an additional purchase, which was originally not planned by the user.

Status Quo Bias. The status quo bias describes the strong tendency of individuals to remain with the status quo as the disadvantages of leaving the current state loom larger than the advantages associated with a change. Kahneman et al. see the status quo bias as a manifestation of an asymmetry of value called loss aversion, that is, "the disutility of giving up an object is greater than the utility associated with acquiring it"

[83, p. 194]. A prominent example is the Austrian organ donor system, which automatically registers every citizen as an organ donor, while in other countries the opposite is the case. In Austria, individuals need to actively decide against organ donation, which positively influenced the participation [58]. In the digital context, many examples can be found where companies set defaults on their websites, such as insurance options on travel websites or delivery options on e-commerce sites. Another example are online configuration tools for cars (e.g., Tesla.com). The car configurator on the Tesla website is a practical application example for nudging, where a nudge in the form of default settings is implemented. When configuring a model, certain packages and options are chosen by default. This procedure is also applied for software products (e.g., pre-selected installation options).

Social Norms. Social norms influence human behavior and can be described as "rules and standards that are understood by members of a group and that guide and/or constrain social behavior without the force of laws" [84, p. 152]. Social norms emerge from "interaction with others; they may or may not be stated explicitly, and any sanctions for deviating from them come from social networks, not the legal system." [84, p. 152]. Moreover, individuals tend to orient towards the behavior of others, searching for social proof when unable to determine the appropriate mode of behavior in a given situation. An example for the application of social norms in nudging is the "most of us wear seatbelts" campaign in the USA in 2002 and 2003 by the Montana Department of Transportation, which aimed to promote safe driving behavior [85]. Amazon's product recommendation systems exhibit an example for calling upon social proof. On the page of a specific product, a recommendation for further products is given, based on what items were bought by other customers ("Customers Who Bought This Item Also Bought"). The group of other customers set a certain standard or a rule for the purchase of a specific product, which the single customer may follow, taking into account the information possessed by others.

Loss Aversion. The psychological principle of loss aversion assumes that losses and disadvantages have greater impact on preferences than gains and advantages [83]. Price benefits can be used to subsidize environmentally friendly options while taxing less environmentally friendly ones [66]. Examples for nudges on Booking.com can be found on the result page of an applied search for a hotel. There, statements such as "Booked 36 times today", "-45% TODAY!", "8 people are looking right now", or "In high demand!" are implemented to trigger the user to not "lose" the offer she found. By giving information about the popularity or limitation, these statements may shorten the purchase decision.

Anchoring and Adjustment. When individuals lack information, they tend to assess or estimate it by using an individual starting point. This initial starting point is either given by the decision frame or the result of a more or less accurate calculation. Consequently, different starting points result in different estimates and are biased toward the considered starting values. Tversky and Kahneman [86] describe this as anchoring and adjustment. For example, the European Energy Label provides information about the energy class and water consumption as well as energy consumption. These labels are used for home electronics, such as washing machines, televisions, or fridges [42]. The exhibited values provide a reference point (anchor) and may serve for users as a tool for comparison between different choice options. Both online and offline retailers often give different (price) options for a product. Apple, for example, offers the iPhone 6s Plus in three capacity options with different prices. The options are displayed at the same moment, while the lowest and the highest price options serve as anchors. This may lead the user to assess the median option relative to the given reference points (prices) influencing her price perception.

Hyperbolic Discounting. According to the concept of hyperbolic discounting, individuals behave inconsistently in terms of time [87]. They value the present and the near-present stronger than the future. Therefore, individuals prefer options with present effects, even though future effects may be greater or better. Rewards such as direct cash payments, vouchers, or price subsidies may serve as nudges to nudge the user toward the better, yet future, choice or action. These nudges have been implemented to promote healthy activities or discourage unhealthy ones [76]. An example for the application in the online sphere can be observed on the website of Europcar, which uses immediate rewards. The result page of Europcar's rental car search displays the prices, where two prices are given for each result. One price saves 9% on the booking if the customer not only books the car but also pays online. This incentive nudges users toward immediate purchase by providing a financial benefit.

Decoupling. When individuals make a decision, they consider the costs of their choice, but this may not be straightforward. According to Prelec and Lowenstein [88], it is more difficult to evaluate the costs of purchases paid by credit card in contrast to cash, as the payment is decoupled from the consumption. As a result, the perceived costs of the decision decrease. This phenomenon is called decoupling [89]. An approach to overcoming decoupling is the disclosure of costs or effects of decisions. The disclosure of environmental costs with energy use or the full costs of credit cards help individuals to understand future costs in the current decision situation and may help to optimize individuals' choices [18]. Media Markt, Europe's market leading retailer for consumer electronics, offers financing and deferred payment for products on its German website. By this means, the retailer wants to decouple the purchase from the actual payment to lower the decision barrier and make purchase more likely.

Priming. Individuals can be prepared for a situation where a decision takes place. Before the decision is made, specific topics, moods, questions, or information can be introduced, for example, by visualizing the consequences of a decision. An example for priming is the nudge of eliciting intentions, such as "Do you plan to vote?" or "Do you plan to vaccinate your child?", before actions or decisions are taken [18]. Priming can be described as the preparation of individuals for the decision moment by gently leading them to the decision. The priming effect can also overlap with framing and other psychological principles [2]. As a result of our search for illustrative examples of priming in the online domain, we identified the Instagram account of Air France as a tool to prime users for a decision. The exhibited pictures prime the users by visualizing consequences or possible outcomes of a decision – in this case, emotional pictures of travelling and destinations. The pictures may nudge the user toward a specific destination or the decision to travel in general.

Availability Heuristic. Individuals tend to judge probabilities of events based on the ease at which they can be recalled. Easily available and often or regularly occurring events are perceived as more likely than less present events, independent from real probabilities [86]. Media campaigns, for example, can induce the imagination that specific risks are more frequent by exhibiting examples of real cases with fatal outcomes (e.g., deaths caused by smoking, plane crashes). Those visual and frequently displayed cases can alter the judgement of individuals toward vulnerability and increased sensitivity to the specific event [90]. Online banner campaigns are a vivid example of a practical implication of a digital nudge making use of the availability heuristic. In the Google Display Network, advertisers can make campaigns available to users by displaying their campaign on the specific ad spaces. Through tracking the user, they can show the ads repeatedly. In the decision moment, their campaign is at the forefront of their mind, and thus, easily available for the users. This may nudge them toward the option of the advertising firm.

4 Conclusion, Limitations, and Further Research

Given the high proliferation of technology in everyday life, more and more purchases as well as life decisions are made on screens. In digital contexts, users often engage in fast and automated decision making, making them prone to making deficient decisions. Against this background, we presented digital nudging as an effective tool to guide the users' decisions by implementing purposefully designed UI design elements. While nudging has been widely discussed outside the IS and HCI domain, little is known about the psychological mechanisms that underlie digital nudging. To address this research gap, we conducted a systematic literature review and identified twenty psychological effects that were investigated in the physical context and that may be transferred to digital environments. In this paper, we presented nine effects in detail as well as exemplary nudges in the physical and digital spheres.

Our research has several implications for theory and practice. First, by presenting the concept of digital nudging, we aim to encourage both researchers and practitioners to incorporate it into their work leveraging the insights into decision making processes and approaches to alter it. It is our intention to inspire behavioral and design-oriented researchers to conduct further research on the effectiveness of digital nudging and thereby advance this increasingly relevant concept. Moreover, we aim to provide new stimuli to practitioners in private and public organizations to create effective UI that benefit both users and organizations. Second, the identified psychological effects and exemplary nudges contribute to HCI research. While HCI scholars are well aware of human psychology and cognitive science, these new insights enhance the theoretical basis of UI and UX design and can be used in design processes and guidelines. Design-oriented researchers can apply psychological effects and nudges when designing IT artefacts to either leverage or counteract the influence of specific psychological effects. Positioning nudges effectively on UI can increase the usability and UX of IT artefacts. Third, for practitioners, the identified psychological effects and exemplary nudges enhance the understanding of decision making and cognitive heuristics and biases at play. UI designers can use these insights to design nudges, i.e., simple interventions for a specific use context and goal. As digital nudges are small changes to an existing UI, their implementation is relatively fast and cheap. Moreover, interventions designed based on empirically validated theory may be more effective compared to a trial-and-error approach, which is often used in practice. Furthermore, our findings help practitioners to better assess whether implemented choice environments serve the intended purpose or steer the user toward an unintended behavior.

The main limitation of this work is that the examples of digital nudges were chosen based on the authors' observation of the websites. We were not able to assess whether the UI design elements were the result of a deliberate nudging development process. Furthermore, as mentioned in the literature, psychological effects partly overlap. Consequently, some of the illustrated nudging examples also overlap, and thus, the underlying psychological effects cannot be clearly differentiated.

Digital nudging unlocks a plethora of further research opportunities. As stated in the introduction, this paper addresses the second and third steps of the nudging development process. Design-oriented researchers could focus on the later steps by designing, implementing and evaluating the effectiveness of digital nudges through lab or real world experiments. From a behavioral research perspective, it would be valuable to investigate the psychological effects in digital contexts to determine whether they show similar predictable effects as in physical contexts. Moreover, it appears promising to examine the effects of specific digital nudges on individuals' decision making, in different digital contexts (e.g., PC, mobile devices, digital signage), as well as to consider different user characteristics. The results may allow for tailoring digital nudges to individual users by leveraging user data and targeting technologies, depending on their current use context and their characteristics.

References

1. Thaler, R.H., Sunstein, C.R., Balz, J.P.: Choice Architecture. In: Shafir, E. (ed.) The Behavioral Foundations of Public Policy, pp. 428–439, Princeton University Press, Princeton (2013)
2. Thaler, R.H., Sunstein C.: Nudge: Improving Decisions About Health, Wealth, and Happiness. Penguin, USA (2009)
3. Lehmann, B.A., Chapman, G.B., Franssen, F.M., Kok, G., Ruiter, R.A.: Changing the Default to Promote Influenza Vaccination Among Health Care Workers. Vaccine, 34, 1389–1392 (2016)
4. Guthrie, J., Mancino, L., Lin, C.T.J.: Nudging Consumers Toward Better Food Choices: Policy Approaches to Changing Food Consumption Behaviors. Psychology & Marketing, 32, 501–511 (2015)
5. Demarque, C., Charalambides, L., Hilton, D. J., Waroquier, L.: Nudging Sustainable Consumption: The Use of Descriptive Norms to Promote a Minority Behavior in a Realistic Online Shopping Environment. J. of Environmental Psychology, 43, 166–174 (2015)
6. Alemanno, A., Spina, A.: Nudging Legally: On the Checks and Balances of Behavioural Regulation. Int. J. of Constitutional Law, 12, 429–456 (2014)

7. Lehner, M., Mont, O., Heiskanen, E.: Nudging – A Promising Tool for Sustainable Consumption Behaviour? J. of Cleaner Production, 134, 166–177 (2015)
8. Borovoy, A., Roberto, C.A.: Japanese and American Public Health Approaches to Preventing Population Weight Gain: A Role for Paternalism? Social Science & Medicine, 143, 62–70 (2015)
9. Marteau, T.M., Ogilvie, D., Roland M., Suhrcke, M., Kelly M.P.: Judging Nudging: Can Nudging Improve Population Health? BMJ, 342, 263–265 (2011)
10. Benartzi, S., Lehrer, J.: The Smarter Screen: Surprising Ways to Influence and Improve Online Behavior. Portfolio (2015)
11. Weinmann, M., Schneider, C., vom Brocke, J.: Digital Nudging, http://dx.doi.org/10.2139/ssrn.2708250 (Accessed: 28.04.2015)
12. Gunaratne, J., Nov, O.: Informing and Improving Retirement Saving Performance Using Behavioral Economics Theory-driven User Interfaces. In: 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 917–920. ACM, New York (2015)
13. Goffart, K., Schermann, M., Kohl, C., Preißinger, J., Krcmar, H.: Using the Default Option Bias to Influence Decision Making While Driving. Int. J. of Human-Computer Interaction, 32, 39–50 (2016)
14. Colin, C., George, L.: Behavioral Economics: Past, Present, Future. Advances in Behavioral Economics. Princeton University Press, Princeton (2004)
15. Kahneman, D.: Thinking, Fast and Slow. Macmillan, New York (2011)
16. Kahneman, D.: Maps of Bounded Rationality: Psychology for Behavioral Economics. The American Economic Review, 93, 1449–1475 (2003)
17. Lichtenstein, S., Slovic, P. (eds.): The Construction of Preference. Cambridge University Press, New York (2006)
18. Sunstein, C.R.: Nudging: A Very Short Guide. J. of Consumer Policy, 37, 583–588 (2014)
19. Oppermann, R.: User-interface Design. In: Adelsberger, H.H., Collis, B., Pawlowski, J.M. (eds.) Handbook on Information Technologies for Education and Training, pp. 233–248, Springer, Heidelberg (2002)
20. Hassenzahl, M., Tractinsky, N.: User Experience – A Research Agenda. Behaviour & Information Technology, 25, 91–97 (2006)
21. Stone, D., Jarrett, C., Woodroffe, M., Minocha, S.: User Interface Design and Evaluation. Morgan Kaufmann, San Francisco (2005)
22. Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. Pearson (2010)
23. Dzida, W.: Das IFIP-Modell für Benutzerschnittstellen. Office Management, 31, 6–8 (1983)
24. Hurtienne, J.: Cognition in HCI: An Ongoing Story. Human Technology, 5, 12–28 (2009)
25. Proctor, R.W., Vu, K.P.L.: Principles for Designing Interfaces Compatible with Human Information Processing. Int. J. of Human-Computer Interaction, 32, 2–22 (2016)
26. Vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In: 17th European Conference on Information Systems, pp. 2206–2217, ECIS Proceedings (2009)
27. Altmann, S., Traxler, C.: Nudges at the Dentist. European Economic Review, 72, 19–38 (2014)
28. Momsen, K., Stoerk, T.: From Intention to Action: Can Nudges Help Consumers to Choose Renewable Energy? Energy Policy, 74, 376–382 (2014)
29. Carter, E.D.: Making the Blue Zones: Neoliberalism and Nudges in Public Health Promotion. Social Science & Medicine, 133, 374–382 (2015)

30. Cioffi, C.E., Levitsky, D.A., Pacanowski, C.R., Bertz, F.: A Nudge in a Healthy Direction. The Effect of Nutrition Labels on Food Purchasing Behaviors in University Dining Facilities. Appetite, 92, 7–14 (2015)
31. Miller, G.F., Gupta, S., Kropp, J.D., Grogan, K.A., Mathews, A.: The Effects of Pre-ordering and Behavioral Nudges on National School Lunch Program Participants' Food Item Selection. J. of Economic Psychology, 55, 4–16 (2016)
32. Yevseyeva, I., Morisset, C., Turland, J., Coventry, L., Groß, T., Laing, C., van Moorsel, A.: Consumerisation of IT: Mitigating Risky User Actions and Improving Productivity with Nudging. Procedia Technology, 16, 508–517 (2014)
33. Brodrick, R.: Nudge Theory: A Cost-effective Method for Increasing Resuscitation Decision Making in Oncology Inpatients. Radiotherapy and Oncology, 115, 693 (2015)
34. Martin, A., Suhrcke, M., Ogilvie, D.: Financial Incentives to Promote Active Travel: An Evidence Review and Economic Framework. Am. J. of Preventive Medicine, 43, 45–57 (2012)
35. Bull, J.: Loads of Green Washing – Can Behavioural Economics Increase Willingness-to-pay for Efficient Washing Machines in the UK? Energy Policy, 50, 242–252 (2012)
36. Luoto, J., Levine, D., Albert, J., Luby, S.: Nudging to Use: Achieving Safe Water Behaviors in Kenya and Bangladesh. J. of Development Economics, 110, 13–21 (2014)
37. Avineri, E.: On the Use and Potential of Behavioural Economics from the Perspective of Transport and Climate Change. J. of Transport Geography, 24, 512–521 (2012)
38. Kattelmann, K.K., Bredbenner, C.B., White, A.A., Greene, G.W., Hoerr, S.L., Kidd, T., Brown, O. N.: The Effects of Young Adults Eating and Active for Health (YEAH): A Theory-based Web-delivered Intervention. J. of Nutrition Education and Behavior, 46, 27–41 (2014)
39. Czap, N.V., Czap, H.J., Lynne, G.D., Burbach, M.E.: Walk in my Shoes: Nudging for Empathy Conservation. Ecological Economics, 118, 147–158 (2015)
40. Nielsen, H.L.: Curating and Nudging in Virtual CLIL Environments. The EuroCALL Review, 22, 40–46 (2014)
41. Hausman, D.M., Welch, B.: Debate: To Nudge or Not to Nudge. J. of Political Philosophy, 18, 123–136 (2010)
42. Ölander, F., Thøgersen, J.: Informing Versus Nudging in Environmental Policy. J. of Consumer Policy, 37, 341–356 (2014)
43. Selinger, E., Whyte, K.: Is There a Right Way to Nudge? The Practice and Ethics of Choice Architecture. Sociology Compass, 5, 923–935 (2011)
44. Woodend, A., Schölmerich, V., Denktaş, S.: "Nudges" to Prevent Behavioral Risk Factors Associated With Major Depressive Disorder. Am. J. of Public Health, 105, 2318–2321 (2015)
45. Mohan, B., Chandon, P., Riis, J.: Promoting Portion Downsizing by Improving Consumer Response to Percentage Cost vs. Percentage Benefit Offers. NA-Advances in Consumer Research, 41, 147–151 (2013)
46. Smeddinck, U.: Regulieren durch „Anstoßen". Nachhaltiger Konsum durch gemeinwohlverträgliche Gestaltung von Entscheidungssituationen? Die Verwaltung, 44, 375–395 (2011)
47. Thorndike, A.N., Sonnenberg, L., Riis, J., Barraclough, S., Levy, D.E.: A 2-Phase Labeling and Choice Architecture Intervention to Improve Healthy Food and Beverage Choices. Am. J. of Public Health, 102, 527–533 (2012)
48. Levy, D.E., Riis, J., Sonnenberg, L.M., Barraclough, S.J., Thorndike, A.N.: Food Choices of Minority and Low-income Employees: A Cafeteria Intervention. Am. J. of Preventive Medicine, 43, 240–248 (2012)

49. Baskin, E., Gorlin, M., Chance, Z., Novemsky, N., Dhar, R., Huskey, K., Hatzis, M.: Proximity of Snacks to Beverages Increases Food Consumption in the Workplace: A Field Study. Appetite, 103, 244–248 (2016)
50. Johnson, E.J., Shu, S.B., Dellaert, B.G.C., Fox, C., Goldstein, D.G., Häubl, G., Larrick, R.P., Payne, J.W., Peters, E., Schkade, D., Wansink, B., Weber, E.U.: Beyond Nudges: Tools of a Choice Architecture. Marketing Letters, 23, 487–504 (2012)
51. Holland, S.: Public Health Ethics. Polity Press, Cambridge (2015)
52. Kooreman, P., Prast, H.: What Does Behavioral Economics Mean for Policy? Challenges to Savings and Health Policies in the Netherlands. De Economist, 158, 101–122 (2010)
53. Arneson, R.J.: Nudge and Shove. Social Theory and Practice, 41, 668–691 (2015)
54. Thaler, R.H., Sunstein, C.R.: Libertarian Paternalism. The American Economic Review, 93, 175–179 (2003)
55. Qizilbash, M.: Informed Desire and the Ambitions of Libertarian Paternalism. Social Choice and Welfare, 38, 647–658 (2012)
56. Kallbekken, S., Sælen, H.: 'Nudging' Hotel Guests to Reduce Food Waste as a Win–win Environmental Measure. Economics Letters, 119, 325–327 (2013)
57. Libotte, E., Siegrist, M., Bucher, T.: The Influence of Plate Size on Meal Composition. Literature Review and Experiment. Appetite, 82, 91–96 (2014)
58. Rebonato, R.: A Critical Assessment of Libertarian Paternalism. J. of Consumer Policy, 37, 357–396 (2014)
59. Monteleone, S.: Addressing the Failure of Informed Consent in Online Data Protection: Learning the Lessons from Behaviour-Aware Regulation. Syracuse J. of International Law and Commerce, 43, 69–191 (2015)
60. Desai, A.C.: Libertarian Paternalism, Externalities, and the "Spirit of Liberty": How Thaler and Sunstein are Nudging us Toward an "Overlapping Consensus". Law & Social Inquiry, 3, 263–295 (2011)
61. Laskowski, M.: Nudging Towards Vaccination: A Behavioral Law and Economics Approach to Childhood Immunization Policy. Texas Law Review, 94, 601–628 (2015)
62. Mols, F., Haslam, S.A., Jetten, J., Steffens, N.K.: Why a Nudge is not Enough: A Social Identity Critique of Governance by Stealth. Eur. J. of Political Research, 54, 81–98 (2015)
63. Van Dalen, H.P., Henkens, K.: Comparing the Effects of Defaults in Organ Donation Systems. Social Science & Medicine, 106, 137–142 (2014)
64. Dolan, P., Hallsworth, M., Halpern, D., King, D., Metcalfe, R., Vlaev, I.: Influencing Behaviour: The Mindspace Way. J. of Economic Psychology, 33, 264–277 (2012)
65. Forwood, S.E., Ahern, A.L., Marteau, T.M., Jebb, S.A.: Offering Within-Category Food Swaps to Reduce Energy Density of Food Purchases: A Study Using an Experimental Online Supermarket. Int. J. of Behavioral Nutrition and Physical Activity, 12, 1 (2015)
66. Hilton, D., Charalambides, L., Demarque, C., Waroquier, L., Raux, C. : A Tax Can Nudge: The Impact of an Environmentally Motivated Bonus/Malus Fiscal System on Transport Preferences. J. of Economic Psychology, 42, 17–27 (2014)
67. Thomas, J.M., Liu, J., Robinson, E.L., Aveyard, P., Herman, C.P., Higgs, S.: Descriptive and Liking Social Norm Messages Enhance the Consumption of a Cruciferous Vegetable in Healthy Students: Sustained Effects After a 24 Hour Delay. Appetite, 101, 218 (2016)
68. Aldrovandi, S., Brown, G.D., Wood, A.M.: Social Norms and Rank-based Nudging: Changing Willingness to Pay for Healthy Food. J. of Experimental Psychology: Applied, 21, 242–254 (2015)
69. Schnellenbach, J.: Nudges and Norms: On the Political Economy of Soft Paternalism. Eur. J. of Political Economy, 28, 266–277 (2012)

70. White, M.D.: The Manipulation of Choice: Ethics and Libertarian Paternalism. Palgrave Macmillan, New York (2013)
71. Korobkin, R.B.: Libertarian Welfarism. California Law Review, 97, 1651–1686 (2009)
72. Jones, L.E., Loibl, C., Tennyson, S.: Effects of Informational Nudges on Consumer Debt Repayment Behaviors. J. of Economic Psychology, 51, 16–33 (2015)
73. Tasoff, J., Letzler, R.: Everyone Believes in Redemption: Nudges and Overoptimism in Costly Task Completion. J. of Economic Behavior & Organization, 107, 107–122 (2014)
74. Bertsimas, D., O'Hair, A.: Learning Preferences under Noise and Loss Aversion: An Optimization Approach. Operations Research, 61, 1190–1199 (2013)
75. Jung, D., Jeong, W. Nudge: A Tool for Better Policy Impacts and its Limitations under Various Policy Contexts. Public Administration Review, 71, 653–656 (2011)
76. Grand, J.L.: The Giants of Excess: A Challenge to the Nation's Health. J. of the Royal Statistical Society: Series A (Statistics in Society), 171, 843–856 (2008)
77. Ratner, R.K., Soman, D., Zauberman, G., Ariely, D., Carmon, Z., Keller, P.A., Kim, B.K., Lin, F., Malkoc, S., Small, D.A., Wertenbroch, K.: How Behavioral Decision Research can Enhance Consumer Welfare: From Freedom of Choice to Paternalistic Intervention. Marketing Letters, 19, 383–397 (2008)
78. Bockstedt, J.C., Goh, K.H., Ng, S.: Contrast Effects in Online Auctions. Electronic Commerce Research and Applications, 12, 139–151 (2013)
79. Lades, L.K.: Impulsive Consumption and Reflexive Thought: Nudging Ethical Consumer Behavior. J. of Economic Psychology, 41, 114–128 (2014)
80. Jacobsen, G.D.: Consumers, Experts, and Online Product Evaluations: Evidence from the Brewing Industry. J. of Public Economics, 126, 114–123 (2015)
81. Johnson, E.J.: Choice Theories: What are They Good for? J. of Consumer Psychology, 23, 154–157 (2013)
82. Kahneman, D., & Tversky, A.: The Simulation Heuristic. In: Kahneman, D., Slovic, P., Tversky, A. (eds.) Judgment Under Uncertainty: Heuristics and Biases, pp. 201–208, Cambridge University Press, New York (1982)
83. Kahneman, D., Knetsch, J.L., Thaler, R.H.: Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. The J. of Economic Perspectives, 5, 193–206 (1991)
84. Cialdini, R. B., Trost, M.R.: Social Influence: Social Norms, Conformity, and Compliance. In: Gilbert, D.T., Fiske, S.T., Lindzey, G. (eds.) The Handbook of Social Psychology, vol. 2, pp. 151–192. McGraw-Hill, New York (1998)
85. Center for Health & Safety Culture Montana State University, http://chsculture.org/mou_projects/most-of-us-wear-seatbelts-campaign-2002-2003/ (Accessed: 26.08.2016)
86. Tversky, A., Kahneman, D.: Judgment Under Uncertainty: Heuristics and Biases. Science, 185, 1124–1131 (1974)
87. Thaler, R.H., Benartzi, S.: Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving. J. of Political Economy, 112, 164–187 (2004)
88. Prelec, D., Loewenstein, G.: The Red and the Black: Mental Accounting of Savings and Debt. Marketing Science, 17, 4–28 (1998)
89. Thaler, R.H.: Mental Accounting Matters. J. of Behavioral Decision Making, 12, 183–206 (1999)
90. Stewart, N., Chater, N., Brown, G.D.: Decision by Sampling. Cognitive Psychology, 53, 1–26 (2006)

Terrorbekämpfung mithilfe sozialer Medien – ein explorativer Einblick am Beispiel von Twitter

Christian Reuter, Katja Pätsch, und Elena Runft

Universität Siegen, Institut für Wirtschaftsinformatik, Siegen, Deutschland
christian.reuter@uni-siegen.de

Abstract. Das Internet und insbesondere soziale Medien werden bekanntermaßen nicht nur zu vermeintlich guten Zwecken genutzt. So findet die Rekrutierung neuer Mitglieder und die Verbreitung von Ideologien des Terrorismus ebenfalls über dieses Medium statt. Aber auch die Terrorismusbekämpfung bedient sich gleicher Werkzeuge. Die Art und Weise dieser Gegenmaßnahmen sowie die Vorgehensweisen sollen in diesem Artikel thematisiert werden. Im ersten Teil wird der Forschungsstand zusammengefasst. Der zweite Teil stellt eine explorative empirische Studie der Terrorismusbekämpfung in sozialen Medien, insbesondere in Twitter, dar. Verschiedene, möglichst charakteristische Formen werden in diesem Rahmen am Beispiel von Twitter strukturiert. Ziel ist es, sich diesem hochrelevanten Gebiet mit dem Ziel von Frieden und Sicherheit aus Perspektive der Wirtschaftsinformatik zu nähern und weiteren Forschungsarbeiten in diesem Gebiet als Grundlage und Ausgangspunkt dienen zu können.

Keywords: Soziale Medien, Twitter, Terrorismus, Krisenmanagement

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Reuter, C.; Pätsch, K.; Runft, E. (2017): Terrorbekämpfung mithilfe sozialer Medien - ein explorativer Einblick am Beispiel von Twitter, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 649-663

1 Einleitung

Die Anschläge in Paris 2015 und in Belgien 2016, um zwei aktuelle Beispiele zu nennen, forderten nicht nur viele Opfer, sondern hatten auch weitreichende politische Folgen. Der sich zu den Anschlägen bekennende sogenannte Islamische Staat (IS) nutzt insbesondere soziale Medien auf eine professionale Art und Weise zur Rekrutierung neuer Kämpfer und zur Werbung für ihre Ideologie [1]. Aber auch aktive Gegner dieses Terrorismus nutzen die gleichen Kanäle zur Terrorbekämpfung. Soziale Medien spielen auf beiden Seiten eine große Rolle. Deren Nutzung durch Unternehmen [2], aber auch durch terroristische Organisationen wurde in der Forschung bereits betrachtet (siehe Kapitel 2). Die Terrorbekämpfung im Web 2.0 hat allerdings wissenschaftlich noch vergleichsweise weniger Beachtung gefunden. Obwohl beispielsweise das Hacker Netzwerk *Anonymous* dem IS den virtuellen Krieg androhte, wurden die daraus resultierenden Folgen wenig in den Medien rezipiert.

Wir fragen uns, wie wird auf die Propaganda der Terroristen geantwortet? Twitter eignet sich nicht zuletzt aufgrund der Zugriffsmöglichkeiten gut, um eine derartige Analyse durchzuführen. Deshalb untersuchen wir in dieser Arbeit, von welcher kompakte Auszüge bereits als Poster-Beitrag veröffentlicht wurden [3], Tweets und Twitter-Accounts im Hinblick auf Anti-Terrormaßnahmen. Hierzu werden wir zunächst den Stand der Forschung zu Terrorismus und sozialen Medien darlegen (Kap. 2). Anschließend erläutern wir unsere Methodik, gefolgt von der qualitativen Analyse der Inhalte (Kap. 3). Schließlich diskutieren wir unsere Ergebnisse (Kap. 4), um ein Fazit aus unseren Erkenntnissen ziehen zu können (Kap. 5).

2 Stand der Forschung: Terrorismus, Propaganda und deren Bekämpfung in sozialen Medien

Dieses Kapitel beschäftigt sich mit dem aktuellen Stand der Forschung bezüglich der Relevanz der sozialen Medien im Zusammenhang mit terroristischer Propaganda sowie existierenden Forschungsergebnissen zu möglichen Gegenmaßnahmen.

Verbreitung terroristischer Propaganda in sozialen Medien. In vielen Arbeiten zu Terrororganisationen und sozialen Medien werden allgemein terroristische Organisationen, aber meist konkret der sogenannte Islamische Staat (IS, ISIS, ISIL, DEASH) thematisiert. So versuchen auch wir allgemein von allen Terrororganisationen zu sprechen, spezifizieren uns aber auch auf den IS, da die aktuelle Situation dies einfordert. Medien, unter anderem die TV-Berichterstattung, spielen im Terrorismus eine große Rolle: „Einen Teil der Verantwortung für die zu beobachtenden Panikreaktionen der Menschen tragen die Medien“ [4]. Einer US-repräsentativen Studie gemäß fördert Angst Vorsorgemaßnahmen; Zorn verhindert sie [5]. Christoph [6] argumentiert: „Ohne ein Bekennerschreiben, ein Abschiedsvideo des Attentäters oder ein letztes Posting im sozialen Netzwerk wäre ein Bombenanschlag nichts als ein Kapitalverbrechen. Durch die terroristische Kommunikationsstrategie wird das Verbrechen erst zum terroristischen Akt“. Terroristen sind jedoch „nicht mehr auf Medienmacher angewiesen, sie sind selbst zum Agens in diesem Spiel geworden“ [6]. Und das nicht ohne Grund: „Terrorismus kann […] nur dann Bedeutung besitzen, wenn er auch auf einer medialen Ebene […] sinnhaft gemacht wird“ [6]. Soziale Medien bieten demnach „den Vorteil der Immersion, das heißt der Verschmelzung von Medium und Botschaft. Die Glaubwürdigkeit terroristischer Narrationen wird so gestärkt, indem sie über vermeintlich seriöse Portale wie YouTube verbreitet werden“ [6]. So entsteht das Gefühl, „der Terrorist sei *einer von uns*, indem er die selben Kommunikationskanäle (‚-waffen‘) nutzt wie der Rezipient selbst“ [6].

Doch nicht nur YouTube dient der Verbreitung: In den vergangenen Jahren ist Twitter zur beliebtesten Internetplattform von Terroristen geworden [7]. Neer und O’Toole [8] haben die Nutzung sozialer Medien durch ISIS untersucht und betonen, dass soziale Medien (insbesondere Twitter) als strategisches Werkzeug genutzt wird, um junge Dschihadisten, Ba'ath Beamte und Frauen für ihre Gewaltüberzeugungen zu begeistern. Klausen et al. [9] betonen, dass die britische Terroristengruppe al-Muhajiroun ihr internationales Netzwerk von YouTube-Kanälen für Propaganda und die Darstellung von gewalthaltigen Inhalten sehr durchdacht ausnutzt. Weinmann und Jost [10] erläutern die Nutzung von Facebook, Twitter und YouTube durch Terrororganisationen zur Rekrutierung und Propaganda: Soziale Medien erleichtern es, Gleichgesinnte zu finden und deren Inhalte zu konsumieren. Gleichzeitig können Terroristen in sozialen Medien auch eine quasi endlose Menge potentieller Mitglieder adressieren, die sonst nicht den Weg in die geschlossenen Foren finden würden. Gemäß der Studie waren zu Beginn Foren die wichtigsten Kontaktpunkte für Mitglieder, Interessenten und Neueinsteiger. Deren Nutzung wurde durch polizeiliche Überwachung erschwert, ebenso gerieten sie in das Fadenkreuz von Nachrichtendiensten, die die Seiten angriffen und stilllegten. Daraufhin wendeten sich immer mehr internetaffine Terroristen den sozialen Netzwerken zu [10]. Dennoch existieren die Foren weiterhin parallel, denn dort debattiert der harte Kern der Community „generelle Entwicklungen in der Szene, führt Diskussionen über ideologische und theologische Fragen und konsumiert Autopropaganda“ [10].

Mehr um den Inhalt geht es bei Zelin [11]: Alle Medien des IS sind auf Arabisch veröffentlicht und nur ein kleiner Anteil ist in andere Sprachen übersetzt. Nach Archetti [12] trägt Online-Propaganda aber nicht nur zur Rekrutierung bei: Ein Dschihad-Video anzusehen könnte sogar die Entschlossenheit von Terrorgegnern erhöhen – folglich eine entgegengesetzte Wirkung. Trotz dieses Zwiespalts gilt: Jede Aufmerksamkeit ist essenziell. Während manche mit Abneigung auf die Videos reagieren, lassen sich andere davon inspirieren – und genau das will der IS erreichen [13].

Die Propagandaziele, in deren Verbreitung sozialen Medien eine Schlüsselkompetenz zugeschrieben wird, reichen von externen Absichten wie der Verbreitung von Angst, bis zu internen wie der Schaffung emotionaler Verbundenheit mit der Zielgruppe [14]. Die Strategie der IS-Mitglieder ist es, durch Posts potenzielle Anhänger aus der ganzen Welt anzuwerben und globale Terrorzellen zu schaffen. Greene [15] thematisiert die taktischen Vorgehensweisen*:* Eine beliebte Taktik zum Verbreiten von ISIS-Propaganda über soziale Medien sind „Twitter-Bomben“, die die beliebtesten Hashtags der Woche nutzen und sie in IS-verwandten Beiträgen verwenden, wodurch ein größeres Publikum erreicht wird [15]. Damit erreicht der IS Personen, die für gewöhnlich nicht an Propagandamaterial interessiert sind und nur zufällig auf die terroristischen Tweets stößt. Ein weiterer Trick zur schnellen Verbreitung ist das gegenseitige Folgen auf Twitter [16]. Meinungsfreiheit, liberale oder unzureichende Inhaltskontrolle auf Twitter begünstigen terroristische Absichten.

Bekämpfung terroristischer Propaganda in sozialen Medien. Einen neuen Blickwinkel auf terroristische Aktionen im Internet eröffnet Gartenstein-Ross [17]: Er gesteht dem IS die erfolgreiche Nutzung von beispielsweise Twitter zu, aber macht gleichzeitig darauf aufmerksam, dass der IS völlig auf den Erfolg dieser Propaganda angewiesen ist und daher eine Schwächung der IS-Kommunikationsstrategie angepeilt werden sollte. Seine Lösungsansätze sind hierfür das Gründen einer kleinen und flinken Einheit speziell zur Widerlegung IS geladener Propaganda [17]. Denn in puncto Glaubwürdigkeit sieht Gartenstein-Ross eine Schwachstelle: Ein Grund, warum IS Nachrichten verwundbar sind, ist, dass Teile davon nicht wahr sind und dadurch der IS beträchtlichen Schaden bezüglich der Wahrnehmung seiner Glaubwürdigkeit riskiert [17]. Ein weiterer Aspekt ist die Professionalität im Umgang mit sozialen Medien. Die sprachlichen Fähigkeiten der Mitglieder (die Statements und Videos in europäische Sprachen übersetzen) [18] tragen zur Verständniserleichterung bei.

Mit generellen Maßnahmen und Strategien zur Terrorbekämpfung beschäftigt sich Turk [19]. Demnach sind die Vereinigten Staaten die wichtigste Weltmacht und Marktführer in der Technologie der Anti-Terror-Entwicklung. Dagegen fokussieren sich Jeberson und Sharma [20] auf die gezielte Ermittlung möglicher Methoden zur Identifizierung Terrorverdächtiger in sozialen Netzwerken. Cheong und Lee [21] beschreiben, dass in Verbindung mit intelligenten Data-Mining-, Visualisierungs- und Filtermethoden diese Daten in einer Wissensbasis gesammelt werden könnten, die für Entscheidungsträger und Behörden von großem Nutzen für schnelle Reaktion und Überwachung während eines solchen Szenarios werden könnten. Sutton et al. [22] thematisieren darüber hinaus den Einsatz von *Backchannels* als einer besonderen Form des Data-Minings zur Informationsgewinnung. Weinmann und Jost [10] erläutern, dass „die Analyse terroristischer Onlinekommunikation, wie sie auf den entsprechenden Social Media-Seiten quasi offen einsehbar ist, [kann] uns viel über die Denkweise von Terroristen, ihre Motivation, ihre Pläne und Ängste verraten“ könne. Statt einer strengeren Zensur radikaler Inhalte sollen demnach „terroristische Kommunikationsstrategien durch einen Mix aus technischen (Hacking) und vor allem psychologischen (Gegenpropaganda) Mitteln“ [10] gestört werden. Hussain und Saltman [23] betonen, dass die generelle Zensur sogar kontraproduktiv sein kann und regen zu positiven Maßnahmen wie den Ausbau von Inhalten gegen den Extremismus an. Ebenfalls folgert Gartenstein-Ross [17], dass die Schwächung der strategischen Kommunikationskampagne des IS ein bedeutsamer Sieg wäre. Aber (glaubwürdige) Gegenpropaganda stammt nicht nur aus dem Ausland: Unter dem Stichwort „Anti IS Humor“ erläutert Al-Rawi [24], dass nach der Veröffentlichung eines ISIS-Videos mit religiösem Gesang Hunderte von arabischen YouTubern begannen, es in lustige Tanzclips zu verwandeln. Darüber ist eine Fokussierung auf Präventivmaßnahmen, auch in Kombination mit (offline) Aufkläraktionen in Schulen, Universitäten oder Gefängnissen, möglich [25].

Die bisherigen Erkenntnisse zur Gegenpropaganda liefern einen Ausgangspunkt für unsere Analyse. Wir explorieren Gegenpropaganda zur Terrorismusbekämpfung und versuchen diese zu systematisieren. In vielen Arbeiten liegt der Fokus auf der Rekrutierung neuer Mitglieder. Bilder und Videos sollen daher die jugendlichen Nutzer von Twitter durch ihre schnelle und einfache Konsumierung beeinflussen. Wir untersuchen und analysieren im Folgenden Gegenmaßnahmen am Beispiel von Twitter.

3 Explorative Studie der Terrorismusbekämpfung in Twitter

Im folgenden Kapitel wird zunächst die verwendete Methodik der Studie beschrieben, um anschließend die resultierenden Kategorien (I) Aufklärung, (II) Parodie/Satire und (III) Hacking hinsichtlich markanter Phänomene zu analysieren.

Methodik. Unsere Analyse fokussiert auf Gegenpropaganda in Tweets und vergleichen Aussagen der Terroristen mit denen von beispielsweise der US-Regierung oder den Medienberichten. Für die Analyse wurde eine qualitative Inhaltsanalyse angewendet. Hierfür war die Verkleinerung des Datenmaterials von Twitter nötig, sodass die Menge qualitativ handhabbar wurde [26]. Auf Twitter finden sich zahlreiche Accounts und Hashtags, die sich dem Kampf gegen den Terror widmeten.

Begonnen wurde die Recherche auf dem Account der US-Regierung @ThinkAgain_DOS, welche versucht, die Propaganda der terroristischen Organisationen aufzudecken. Von diesem Account wurden viele Hashtags wie #FreeTheWorld oder #VivaSAA gefunden, die der Terrorismusbekämpfung dienen. Dazu gehören #ThinkAgainTurnAway, #Daesh, #DaeshLiesExposed, #alleyesonISIS, #IstandwithSyria, #FreeTheWorld, #Syria, #ISIS, #trollingday, #trollingisis, #isismovies, #ISIL, #NotInOurName, #No2ISIS, #OpIceISIS und viele mehr. Außerdem konnten über die Twitter-Funktionalität (Vorschlag) mehrere User mit ähnlichen Interessen gefunden werden, z.B. @OYEQF, @QuilliamF, @Sil3nceBl4ck, @AnonymousMedia, @AntiTerrorismTR, @Active_Change_F. Das in jenen Accounts und Hashtags vom 1. bis 31. März 2016 identifizierbare Material wurde gesammelt und gespeichert.

Account/Hashtag	**Intention**	**Methode**
@ThinkAgain_DOS	Aufklärung	Verbreitung und Kommentierung aktueller Veröffentlichungen zum Thema
@AverageMohamed	Aufklärung	Informationsverbreitung mithilfe von Cartoons
@Operation_ISIL	Hacking	Aufruf durch Hashtags, auffällige Accounts zu melden
@isis_karaoke	Parodie / Satire	Posten von Fotos von IS-Kämpfern mit Zeilen aus bekannten Pop-Songs unter dem Motto „dropping songs, not bombs“
#TrollingISIS bzw. #TrollingDay	Parodie / Satire	Posten modifizierter, verhöhnender ISIS Bilder
#ISISChan bzw. ISIS_Chan	Parodie / Satire	Google-Bombe: Ergebnisse der Bildersuche zu „ISIS“ mit Darstellungen des Anime-Mädchens mit grünen Haaren im schwarzen Dschihadisten-Outfit mit einer Melone
#ISISCrappyCollageGrandPrix	Parodie / Satire	japanische Photoshop und Hashtag Kampagne zur Verhöhnung von IS-Kämpfern

Tabelle 1: Übersicht über ausgewählte Accounts und Hashtags, deren Intention und Vorgehen

Wir verdichteten das gefilterte Material, indem mithilfe von Open Coding [27] Kategorien erstellt wurden, aus denen wir schließlich spezifische Beispiele auswählten und exemplarisch analysierten. Eine Limitation ist die damit nicht erreichte Repräsentativität; diese war jedoch auch nicht das Ziel dieser Arbeit. Es geht hier nicht um die Bestätigung von vorab festgelegten Hypothesen, sondern um die Entwicklung von Hypothesen und Theorien aus dem Material [27]. Weiterhin können Bias-Effekte aufgrund des explorativen Vorgehens mit Identifikation markanter Merkmale nicht ausgeschlossen werden. Überdies wurden englische Tweets untersucht – somit fokussiert die Arbeit Wirkungen und Aktivitäten im englischsprachigen Raum.

Insgesamt ließen sich die gefundenen Posts, Hashtags und Accounts drei (nicht immer vollständig trennscharfen) Kategorien zuordnen: (I) Aufklärung, (II) Parodie/Satire und (III) Hacking. Diese basieren auf den im Codierungsprozess entstandenen und später verdichteten Unterkategorien (Presse, Bilder, Trolling, Anime, Grand Prix, Verulkung). Im Folgenden sollen die Hauptkategorien untersucht und jeweils exemplarische Beispiele daraus analysiert werden. Wir versuchen aus der Masse markante Phänomene herauszuarbeiten, was jedoch nicht immer gelingt. Beispielsweise sind in der Kategorie des Hackings Accounts und Trends schnelllebig und austauschbar. Sie ähneln sich so stark, dass kaum markante Unterschiede festgestellt werden können.

Kategorie I: Aufklärung. Unter Aufklärung verstehen wir die Versuche, der terroristischen Propaganda mit Logik zu antworten, um sie so außer Kraft zu setzten. Laut Duden ist es eine „völlige Klärung“ im Sinne einer „Darlegung, die über bisher unbekannte Zusammenhänge aufklärt“. Interessant sind die beiden Accounts @ThinkAgain_DOS und @AverageMohamed, da sie das gleiche Ziel haben, aber sehr verschieden arbeiten. @ThinkAgain_DOS sollte als offizieller US-Regierungs-Account zur Terrorbekämpfung auf Twitter eine Vorreiterrolle einnehmen. @AverageMohamed dagegen bekämpft den Terror, indem er versucht, mit Cartoons Informationen zu verbreiten. Die praktische Zielgruppe, die es zu untersuchen gilt, sind demnach die Accounts @ThinkAgain_DOS und @AverageMohamed, deren Aktivitäten der Aufklärung der Allgemeinheit sowie potenziellen IS-Anhängern gelten.

Abbildung 1: Tweet von @ThinkAgain_DOS

Abbildung 2: Tweet von @syrmukhabarat

Eine bekannte Methode stellt die Verbreitung und Kommentierung aktueller Veröffentlichungen zum Thema dar. Im @ThinkAgain_DOS Tweet (02.03.2016) verbreiten und kommentieren sie einen Artikel der Zeitung *The Independent* (Abbildung 1), der einen weiterführenden Link zu einem Bericht über britische Imame enthält, die sich im Irak über den Kampf gegen die ISIS informieren. Dies geschieht in der Hoffnung, dass die Imame ihre Eindrücke aus dem Irak zu Hause verbreiten und so der weiteren Radikalisierung entgegenwirken können. Wir sehen, dass die Maßnahme, auf die gesetzt wird, die Aufklärung über die tatsächlichen Gegebenheiten in IS

Gebieten innerhalb der sozialen Medien ist. @Selbyist hat auf den Tweet geantwortet: „*This will need thorough investigation. I don't trust that rag Independent, and who's paying for this, George fucking Soros?*” (02.03.2016). Dadurch wird zum einen „Independent“ als Quelle, der er/sie nicht vertraut, kritisiert und nach der Finanzierung gefragt. Hier erkennt man, dass Aufklärung nur über Medienberichte zur Glaubwürdigkeit nicht ausreicht.

Die User sind kritisch, vor allem, wenn es von der US-Regierung stammt, der kein Vertrauen entgegengebracht wird. Exemplarisch belegbar ist dies mit einer Diskussion: @syrmukhabarat postete eine Collage aus sechs kleineren Bildern von männlichen, blutüberströmten Leichen, die mit dem Logo der syrisch-arabischen Armee dekoriert ist, und einen Link zu einem Zeitungsartikel (Abbildung 2): „*Here are the Rebels that were ambushed trying to escape Krak des Chevaliers into Lebanon*“ (20.03.2014). Der Kommentar von @syrmukhabarat ordnet die Toten den Rebellengruppen zu. Der Zeitungsartikel titelt: „*11 rebels killed fleeing famed Crusader fort: Syria army*“ (20.03.2014). Damit stellt er zwischen beiden einen Zusammenhang her, obwohl die Quelle des Bildes unklar ist. @ThinkAgain_DOS kommentiert: *„[…] They could also be #alquaeda fighters whom #Assad used to send to Iraq and Lebanon #thinkagainturnaway”* (20.03.2014). Dadurch wird die Diskussion über die nicht vorhandene Quelle des Bildes eröffnet und weitere Spekulationen angestellt.

Abbildung 3: Bildkampagne von @ThinkAgain_DOS

Abbildung 4: Bildkampagne von @ThinkAgain_DOS

Abbildung 5: Tweet von @AverageMohamed

Neben den Presseartikeln postet @ThinkAgain_DOS auch Bilder oder Videos, die aufklärerische Botschaften übermitteln sollen. Unter #UnitedAgainstDaesh wurde ein Bild (Abbildung 3, 09.03.2016) gepostet, welches eine irakische Frau und ein Mädchen zeigt, die nebeneinander ernst auf einem Sofa sitzen: *„'For one hour a day they electrocuted me: cables to my head, hands and feet. I was crying and begging him to stop, but he wouldn't listen.' Iraqi woman held captive by ISIS for 4 months"*. Der Kommentar zu dem Bild lautet: *„#UnitedAgainstDaesh: Coalition seeks to destroy the evil perpetrators of extreme violence against"*. Ziel dieses Posts ist es, den Vorwürfen gegen ISIS, gewalttätig und skrupellos zu sein, ein Gesicht zu verleihen.

Am 13.03.2016 postet @ThinkAgain_DOS ein weiteres Bild (Abbildung 4), auf dem eine Frau und ein Mädchen abgebildet sind, welche gemeinsam ein Schild hochhalten. Der Text darauf lautet: *„ISIS DOES NOT REPRESENT ISLAM“*. Frau und Mädchen blicken genau in die Kamera. Dies erweckt den Eindruck, dass sie den Betrachter des Bildes direkt ansehen. Dieses vermutliche Mutter-Tochter-Gespann soll jedoch kein Mitleid erregen, sondern dem Betrachtet vermittelt, dass hier zwei starke muslimische Frauen für die Aufklärung über ISIS einstehen. Das Bild bleibt unkommentiert, es wurde nur mit den Hashtags #NoToDaesh und #DaeshLiesExposed versehen. Die beiden Frauen stehen symbolisch für alle modernen Musliminnen, die ISIS nicht unterstützen und sich nicht diskriminieren lassen.

@AverageMohamed, eine Art Kunstfigur, hingegen produziert Cartoons, die zum Ziel haben, Jugendliche durch Aufklärung vor Radikalisierung zu schützen. Über ihn und seine Arbeit wurde in *The Guardian, The Telegraph* sowie in Radio *Bayern 2* berichtet. Er postet einen Link zu einem seiner YouTube-Videos (Abbildung 5, 21.10.2015), welches mit „Identity in Islam“ betitelt und mit verschiedenen Hashtags (#almuhajirah, #jondi) versehen ist. Diese Hashtags werden unter anderem auch von terroristischen Organisationen genutzt und sollen die Propaganda-Arbeit unterlaufen und radikalisierungsgefährdete Jugendliche erreichen. Das Video behandelt die Frage nach Gesellschaft, Religion und Identität. Er wirbt für Vielfalt und Toleranz und erklärt diese zu Bestandteilen des Islams. Damit versucht er Argumenten der Extremisten entgegenzuwirken, die behaupten, es sei Allahs Wille alle Ungläubigen zu töten.

Abbildung 6: #TrollingDay IS-Kämpfer als Quietscheenten

Abbildung 7: #TrollingDay IS-Kämpfer mit Ziegen

Kategorie II: Parodie und Satire. Parodie ist laut Duden eine komisch-satirische Nachahmung durch Stilmittel wie Verzerrung und Übertreibung. Die Satire ist eine Kunstgattung, die an Ereignissen Kritik übt und sie der Lächerlichkeit preisgibt. Beides zielt darauf ab, Spott über ernsthafte Begebenheiten zu äußern.

Unter dem Hashtag #TrollingISIS bzw. #TrollingDay konnten wir hauptsächlich Bilder identifizieren, die sich in offensichtlicher Art über den IS lustig machen. Das Hacker-Netzwerk *Anonymous* erklärte den 11. Dezember zum „ISIS Trolling Day“ und rief auf der Webseite Ghostbin dazu auf, die satirische Stärke des Internets zu nutzen, um Memes und Cartoons zu posten, die den IS ins Lächerliche ziehen. Dieser Plan ist Teil der Operation von *Anonymous*, die seit den Anschlägen am 13.11.2015 in Paris Cyberkrieg gegen ISIS führen. Unter #TrollingISIS und #TrollingDay wurde eine Flut modifizierter ISIS Bilder gepostet. Ein beliebtes Element, das sich schnell verbreitete, war das Ersetzen der Köpfe von den ISIS-Kämpfern durch gelbe Quietscheenten-Köpfe. Statt Waffen halten die Kämpfer Klobürsten in die Luft (Abbildung 6).

Explizit von *Anonymous* aufgerufen wurde zu der Benutzung von Ziegen auf den Bildern: „*post photos of goats while @ing Isis members with captions talking about* *their wives*“. Diese Art der Parodie beruht auf dem Gerücht, der Prophet Mohammed habe Sodomie mit Ziegen gepflegt. Da Geschlechtsverkehr im Islam weitestgehend ein Tabu-Thema ist, trifft dieser Vorwurf ganz besonders den Nerv. Diese Art der Kritik am Islam, insbesondere am IS, ist somit stark politisch motiviert. In Abbildung 7 sind IS-Kämpfer zu sehen, die auf überdimensionalen Ziegen reiten und eine schwarze ISIS-Fahne mit der Aufschrift „We Fuck Goats“ schwingen.

Abbildung 8: #TrollingDay IS-Kämpfer mit Dildos

Abbildung 9: #TrollingDay Abu Bakr al-Baghdadi mit Waschmittel

Eine weitere beliebte Darstellung sind ISIS-Kämpfer in Verbindung mit Homosexualität. Im Islam gilt gleichgeschlechtlicher Verkehr als Unzucht und ist untersagt. In Abbildung 8 sind bewaffnete IS-Kämpfer zu sehen, deren Geschosse in überdimensionale Dildos verwandelt wurden. Ein anderes Thema verschiedener Tweets des Trolling Days ist die mangelnde Intelligenz des IS. Der abgebildete Post zeigt zum Beispiel den ISIS-Anführer Abu Bakr al-Baghdadi mit dem Waschmittel *Daesh* (Abbildung 9). Der Hashtag #Daeshbag ist angelehnt an das Wort „douchebag“, was umgangssprachlich Trottel bedeutet. Mit der Waschmittelwerbung wird der IS – im wahrsten Sinne – verweichlicht dargestellt. Die Gehirnwäsche, die durch Propaganda und Rekrutierung stattfindet, macht die Anhänger zu Marionetten des IS, individuelle Meinungen werden abgelehnt, so die Intension der Gegenpropaganda.

Neben den politisch motivierten Satiremaßnahmen, die eine zielgerichtete Intention der Gegenpropaganda bilden, gibt es auch reine Parodien des IS. Ein solches Beispiel ist das Hashtag #ISISChan oder „ISIS_Chan“. Dort findet man bei Twitter Bilder eines grünhaarigen Anime-Mädchens im schwarzen Dschihadisten-Outfit mit einer Melone. ISIS-Chan wurde von Usern der japanischen Textboard-Seite 2channel kreiert (Abbildung 10). Am 24.01.2015 wurde ein Thread mit dem Titel „*Let’s make ISIS into a moe girl and send it to them!*”[1] erstellt. Eine Google-Bombe sollte die Ergebnisse der Bildersuche zu „ISIS“ mit Darstellungen von dem Anime-Mädchen füllen. In den ersten 24 Stunden wurden die Hashtags auf Twitter 9.000 Mal verwendet (ebd.). Abbildung 11 zeigt ISIS-Chan mit einer Melone auf einem Teller. Sie hält das Messer falsch herum an ihre Kehle. Der Zeichner erklärte, dass sie dem ISIS beibringen wolle, wofür Messer wirklich da sind - nämlich um Melonen zu schneiden.[2]

[1] http://knowyourmeme.com/memes/isis-chan
[2] http://www.dw.com/en/melons-instead-of-tanks-meet-isis-chan-the-anime-that-fights-islamic-state/a-18698611

Abbildung 10: Richtlinien für ISIS-Chan

Abbildung 11: ISIS-Chan mit Melone

Eine japanische Photoshop und Hashtag Kampagne zeigt bearbeitete Bilder des IS mit zwei japanischen Geiseln (#ISISCrappyCollageGrandPrix[3]). Am 20.01.2015 veröffentlichte ISIS ein Video auf YouTube, in dem ein maskierter Kämpfer droht, zwei japanische Geiseln umzubringen, wenn Japan keine zwei Millionen Dollar Lösegeld zahle. Als Antwort twitterten japanische User Parodien des Videos mit dem Hashtag. Abbildung 12 zeigt den IS-Anhänger wie er mit seinem Messer einen Kebab schneidet, worauf teilweise Hass-Kommentare der IS-Anhänger unter den Bildern erschienen, wie: „*Japanese people, You are so optimistic Is it because he said 5800 kms you think you are safe zone. We have army everywhere.*“[4]

Abbildung 12: #ISISCrappyCollageGrandPrix IS-Kämpfer mit Döner

Abbildung 13: @isis_karaoke IS-Kämpfer tanzen vermeintlich zu „Styain‘ alive“

Am 26.12.2015 rief der selbsternannte Kalif des IS Abu Bakr al-Baghdadi in einer Audiobotschaft alle Muslime dazu auf, sich dem Kampf anzuschließen. Allerdings sagten viele Muslime auf Twitter dem IS mit Begründungen wie „*Mein Goldfisch ist krank, ich muss leider zum Tierarzt*“ oder „*Sorry, ich bin beschäftigt mit Netflix*

[3] http://www.stuff.co.nz/world/asia/65395382/japan-launches-crappy-meme-war-with-islamic-state
[4] http://knowyourmeme.com/memes/isis-crappy-collage-grand-prix 21.01.2015

gucken.“ ab. Den Muslimen war es wichtig zu zeigen, dass normale Bürger wichtigere Dinge zu tun haben, als einem weltweiten Religionskrieg beizutreten. Der Account ISIS_Karaoke agiert nach dem Motto „dropping songs, not bombs“ und versehrt Fotos von IS-Kämpfern mit Zeilen aus bekannten Pop-Songs. Hintergrund dazu ist das neuerliche Musikverbot des IS in Racca, der Hauptstadt des sogenannten Kalifats[5]. Abbildung 13 zeigt Dschihadisten, die vermeintlich zu „Stayin‘ alive“ von den Bee Gees performen.

Abbildung 14: Tweet von @Operation_ISIL

Abbildung 15: Tweets zu #GhostSecGroup

Kategorie III: Hacking. Eine weitere Kategorie der Terrorismusbekämpfung umfasst das Hacking. „*Das Hacker-Netzwerk Anonymous hat dem Islamischen Staat (IS) nach der Terrorserie in Paris mit martialischen Worten den virtuellen Krieg erklärt*“[6]. Nach den Pariser Anschlägen 2015 haben sich Hacker, auch außerhalb von *Anonymous*, weltweit dem Kampf gegen ISIS in den sozialen Netzwerken verschrieben.

Abbildung 14 zeigt den Tweet von @Operation_ISIL vom 16.03.2016. Er listet Hashtags auf, welche zu ähnlichen Themen führen und ein Word-Dokument mit einer Reihe von Accounts, die unschädlich gemacht wurden. Abbildung 15 verdeutlicht, dass für jeden die Möglichkeit besteht, sich am Kampf gegen extremistische Gruppen zu beteiligen. Hierfür werden unter entsprechender Anleitung alle User aufgefordert, auffällige Accounts zu melden. Weitere Aktivitäten bleiben dem sozialen Netzwerk verborgen, denn es handelt sich um illegale Aktivitäten, die nicht für die Öffentlichkeit bestimmt sind. Der Erfolg dieser Hackerangriffe bleibt dabei fraglich, da trotz vieler Sperren zahlreiche Seiten aufrufbar sind. Hacker können ebenfalls von ISIS-Usern „gejagt“ werden. Sie bekämpfen sich untereinander und rufen sich gegenseitig zur Jagd auf. In dem Post vom 18.03.2016, macht @intel_ghost auf einen deutschen 'pro-IS' Rapper aufmerksam und fordert die Community auf „Go get him!“. Die Hacker-Szene beteiligt sich rege an den parodistischen Bildern und Videos und verbreitet diese weiter. Hacking bezieht sich somit sowohl auf illegale Aktivitäten, wie beispielsweise die Sperrung von Accounts und den Aufruf an die Bevölkerung, verdächtige Personen zu melden, als auch auf legale Aktivitäten, indem parodistische Medien vermehrt werden.

[5] http://www.stern.de/digital/online/isis-karaoke--satire-auf-twitter-bringt-is-kaempfer-zum-traellern-von-britney-spears-songs-6423078.html
[6] http://www.focus.de/politik/ausland/diese-attacken-koennen-nicht-ungestraft-bleiben-anonymous-erklaert-dem-is-per-videobotschaft-den-krieg_id_5089975.html

4 Diskussion: Wirksamkeit der Gegenmaßnahmen

Soziale Medien können einen Beitrag zur Terrorbekämpfung leisten, wie bereits die Literaturstudie dargestellt hat [10, 17, 20, 21, 24]. Die Maßnahmen gegen terroristische Propaganda auf Twitter sind, wie wir beobachten konnten, sehr rege. Die Hauptarbeit von @ThinkAgain_DOS, welche hier stellvertretend für die US-Regierung eingesetzt ist, besteht nur im Verbreiten von Artikeln. Diese Art der aufklärerischen Maßnahmen ist sehr passiv und reaktionär. Die „Aufklärer" erheben für sich den Anspruch, die Wahrheit zu verbreiten, im Gegensatz zu den Lügen der Terroristen. Gartenstein-Ross [17] führt aus, dass die Schwachstelle des IS seine Glaubwürdigkeit ist und terroristische Kommunikationsstrategien durch technische (Hacking) und vor allem psychologische (Gegenpropaganda) Mittel gestoppt werden können. Allerdings haben die Gegenmaßnahmen das gleiche Problem der Glaubwürdigkeit, wie wir in der Analyse feststellten (vgl. Diskussion ausgelöst von @syrmukhabarat vom 20.03.2014). So haben auch die USA Probleme mit der Glaubwürdigkeit, allerdings zählen sie in den Augen ihrer Gegner selbst zu den Terrorristen. An dieser Stelle haben User wie @AverageMohamed bessere Chancen, positiv wahrgenommen zu werden. Er ist als Muslim glaubwürdiger, wenn er über den Islam aufklärt. Insgesamt werden unterschiedliche Aufklärungsstrategien sichtbar, die auch mit der jeweils eingenommenen Rolle (Regierung, lokale Organisation, politische Organisation, Privatperson) zusammenhängen.

Es fällt überraschenderweise auf, dass die vermeintlich privaten User wesentlich produktiver sind, als offizielle Accounts: In unserer Studie verzeichneten diese eine wesentlich höhere Anzahl an Retweets. Die offiziellen Accounts (z.B. @TheGEC oder @ThinkAgain_DOS) machen sich reichweitenerhöhende Hashtags weniger zunutze als die private Szene. Möglicherweise liegt dies daran, dass von offizieller Seite mehr verdeckt gearbeitet wird. Die Twitter-Nutzer bekämpfen mit Strategien der Gegenpropaganda, wie beispielsweise Satire, den gemeinsamen Feind IS. Parodistische Bildbearbeitungen erfreuen sich großer Beliebtheit im Netz und animieren offenbar zum Mitmachen. Einerseits reagieren Terrorristen mit extremen Drohungen auf die bearbeiteten Bilder und den damit einhergehenden Spott, andererseits gilt auch für den IS, dass jede Beachtung und Verbreitung gut ist. Die aufklärerische Arbeit findet ebenfalls Beachtung, jedoch erhält sie bei Weitem nicht so viel Aufmerksamkeit, wie die satirische Herangehensweise – zumindest im Sinne der von uns gezählten Retweets. Der IS arbeitet sehr viel über Bilder und Videos, daher ist es effektiver auf gleicher Ebene zu antworten, statt mit Texten und Statements.

Algorithmus-basierte Scans könnten „bis zu einem gewissen Grad unterstützen und entlasten, in absehbarer Zukunft kann ein voll-automatisierter Prozess aber keine faire und zensurfreie Überprüfung gewährleisten" [10]. Gerade Bots zur automatischen Meinungsmache stellen eine große Herausforderung dar [28]. Hackerangriffe auf IS-Accounts scheinen alleine wenig erfolgversprechend zu sein, da im schnelllebigen Web anstelle eines gelöschten Accounts sofort zwei neue Accounts entstehen. Dennoch ist diese Arbeit als Zeichen der Gegenwehr von Bedeutung. Sinnvoll, aber nicht einfach, wäre es, gezielt die Grundstruktur terroristischer Online-Auftritte zu infiltrieren und das Problem grundlegender anzugehen, als nur oberflächlich Accounts zu löschen.

Einer terroristischen Gruppierung entgegenzutreten, die nicht einmal der eigene Tod abschrecken kann, ist nur mit genügend Scharfsinn und Galgenhumor zu bewältigen.

Das scheinbar Schlimmste für den IS ist es, in Unwürde oder Lächerlichkeit zu fallen. Viele Reaktionen der IS-Vertreter auf parodistische oder satirische Tweets sind aggressiv und gewaltandrohend. Entsprechende Accounts werden danach oft gehackt oder gemeldet, sodass es problematisch ist, dies im Nachhinein nachzuvollziehen. An diesen Reaktionen ist abzulesen, dass von den Satirikern ein Nerv getroffen wurde. Diese Schwachstelle lässt sich mit Parodie und Satire angreifen, die schnell kollektiven Anklang findet. Allerdings besteht die Möglichkeit, dass der IS daraufhin seine angedrohten Strafen aufgrund der Beleidigung in die Tat umsetzt oder die Weiterverbreitung des Propagandamaterials die Aufmerksamkeit auf den IS lenkt. Eine subtilere Weise sich den Terroristen zu widersetzen, zeigten die Muslime mit ihren Absagen via Twitter zum Beitritts-Aufruf für den heiligen Krieg. Einige User stützten ihre Antworten mit Belegen aus dem Koran, womit sie dem IS die Argumentationsgrundlage entzogen.

5 Zusammenfassung und Fazit

Soziale Medien spielen sowohl im Terrorismus als auch in der Terrorbekämpfung eine Rolle. Auch wenn die Forschung zu sozialen Medien mit sowohl verhaltenswissenschaftlichen, als auch gestaltungsorientierten Studien mittlerweile einen festen Platz in der Wirtschaftsinformatik (WI)-Forschung eingenommen hat, wurde mit Blick auf frühere WI-Artikel zum Terrorismus in sozialen Medien vergleichsweise wenig publiziert. Dieser Beitrag beschäftigt sich mit der Terrorbekämpfung in sozialen Medien. Im ersten Teil wurde der Stand der Forschung im Bereich Terrorismus, Propaganda und deren Bekämpfung in sozialen Medien analysiert und hierauf aufbauend im zweiten Teil die Nutzung von Twitter beispielhaft exploriert.

Betrachtet man die Gesamtheit der auf Twitter aufgetretenen Phänomene, die wir in (I) Aufklärung, (II) Parodie und Satire sowie (III) Hacking klassifiziert haben, wird deutlich, dass eine Maßnahme allein nicht erfolgversprechend ist. Aufklärung ist schlussendlich das Einzige, was vermeintliche Lügen der Terroristen aufzudecken vermag, aber sie muss glaubwürdig kommuniziert werden. Mit Hilfe des Hackings oder der Parodie und Satire erhalten die Twitter-Nutzer die Möglichkeit, als Einheit gegen die Terroristen vorzugehen. So ist es auch möglich, dem Gefühl der Hilflosigkeit angesichts des mächtigen Feindes entgegenzuwirken. Bots, die hier automatisiert arbeiten, sind als große Herausforderung wahrzunehmen. Bei Hacking sind jedoch auch rechtliche und ethische Fragen zu beachten. Die Resonanz für Satire ist da, aber die Gegenmaßnahmen müssen darüber hinausgehen.

Die Maßnahmen im Kontext der Terrorbekämpfung bieten viele Möglichkeiten, die lohnenswert erscheinen, deren Essenz wir kompakt wie folgt darstellen:

(1) Massenbewegungen starten (sich als Einheit gegen den IS stellen)
(2) Authentizität und Glaubwürdigkeit vermitteln (besonders für offizielle Seiten)
(3) Parodie und Satire zur kritischen Reflektion nutzen (besonders für private Stellen)
(4) Gegenwehr auf Augenhöhe (Muslime wenden sich gegen ISIS)
(5) Hacking durch spezialisierte Gruppen (nicht nur oberflächliche Beseitigung)
(6) Aufklärung verständlich vermitteln (Ansprache in Sprache der Bevölkerung)

Selbstverständlich hat diese Studie Limitationen: Dieser explorative Überblick konnte nur einen Teil der Aktivitäten identifizieren. Es wurde mithilfe qualitativer Methoden ein Einblick gewehrt, nicht jedoch das gesamte Gefüge adressiert (siehe Methodik, Kap. 3). In weiteren Studien sollte unter Zuhilfenahme quantitativer Methoden der (sozialen) Netzwerkanalyse sowie größerer Datensätze [29] das Geflecht der Gegenmaßnahmen untersucht werden. Vermutet wird, dass ein wachsender Teil der Propaganda durch Bots erfolgt, sich also der Konflikt mehr auf die technische Ebene verlagert (H1). Zur validen Beurteilung der Wirksamkeit der Gegenmaßnahmen erfordert es eine längerfristige Beobachtung der Reaktionen auf die IS-Propaganda. Es ist auch mithilfe von Social Media Analyse [31] zu untersuchen, ob Parodie und Satire insb. von Sympathisanten dieses Humors konsumiert werden, oder Gefährdete zur kritischen Reflektion anregen (H2). Die Aktivitäten der Opposition, von Aktivisten und Flüchtlingen, die potentiell eine große Rolle spielen, sollten in diesem Rahmen auch untersucht werden [32] (H3). Weiterhin wäre genauer zu untersuchen ob und wie Privatsphäre und Terrorbekämpfung in sozialen Medien durch die Bevölkerung abgewogen werden [33] (H4). Erkenntnisse der Crisis Informatics [30], die sich sonst mehr mit Krisenmanagement beschäftigen, sollten wo möglich auf die Terrorbekämpfung übertragen werden.

Danksagung. Die Forschungsarbeiten wurden im Rahmen des EU-FP7-Projekts „EmerGent“ (Fö.-Kz. 608352) gefördert. Ein Short-Paper/Poster mit Ausschnitten dieser Arbeiten wurde auf der Konferenz Mensch & Computer [3] vorgestellt.

Literaturverzeichnis

1. Seib, P., Janbek, D.M.: Global Terrorism and New Media: The post-Al Qaeda Generation. Routledge, New York (2011).
2. Koch, M., Richter, A.: Enterprise 2.0 - Planung, Einführung und erfolgreicher Einsatz von Social Software in Unternehmen. Oldenbourg-Verlag (2009).
3. Reuter, C., Pätsch, K., Runft, E.: Terrorismus und soziale Medien – Propaganda und Gegenpropaganda. In: Mensch & Computer: Tagungsband. GI., Aachen, Germany (2016).
4. Jackob, N.: Die Diffusion von Terrormeldungen, die Wirkung von Anschlägen auf die öffentliche Meinung und die Folgen für das Vertrauen in der Demokratie. In: Glaab, S. (ed.) Medien und Terrorismus. pp. 155–174. Berliner Wissenschafts-Verlag, Berlin (2007).
5. Lerner, J.S., Gonzalez, R.M., Small, D.A., Fischhoff, B.: Effects of fear and anger on perceived risks of terrorism: A national field experiment. Psychol. Sci. 14, 144–150 (2003).
6. Christoph, S.: Funktionslogik terroristischer Propaganda im bewegten Bild. J. Deradicalization. Fall/15, 145–205 (2015).
7. Khayat, M.: Jihadis’ responses to widespread decline in participation on jihadi forums, increased use of Twitter. MEMRI Inq. Anal. 955, (2013).
8. Neer, T., O’Toole, M.E.: The Violence of the Islamic State of Syria (ISIS): A Behavioral Perspective. Violence Gend. 1, 145–156 (2014).
9. Klausen, J., Barbieri, E.T., Reichlin-Melnick, A., Zelin, A.Y.: The YouTube Jihadists: A Social Network Analysis. Perspect. Terror. 6, 36–53 (2012).
10. Weimann, G., Jost, J.: Neuer Terrorismus und Neue Medien. Zeitschrift für Außen- und Sicherheitspolitik. 8, 369–388 (2015).

11. Zelin, A.Y.: Picture Or It Didn't Happen: A Snapshot of the Islamic State's Official Media Output. Perspect. Terror. 9, 85–97 (2015).
12. Archetti, C.: Terrorism, Communication and New Media: Explaining Radicalization in the Digital Age. Perspect. Terror. 9, 49–59 (2015).
13. Ryan, L.: ISIS Is Better Than Al-Qaeda At Using the Internet, http://www.defenseone.com/technology/2014/10/isis-better-al-qaeda-using-internet/96308/.
14. Chatfield, A.T., Reddick, C.G., Brajawidagda, U.: Tweeting propaganda, radicalization and recruitment. In: Proc. Digital Government Research. pp. 239–249. ACM Press (2015).
15. Greene, K.J.: ISIS: Trends in Terrorist Media and Propaganda. Int. Stud. Capstone Res. Pap. Paper 3, 1–59 (2015).
16. Barrett, R.: The Islamic State. The Soufan Group (2014).
17. Gartenstein-Ross, D.: Social Media in the Next Evolution of Terrorist Recruitment. Hear. before Senat. Comm. Homel. Secur. Gov. Aff. Found. Def. Democr. 1–11 (2015).
18. Gates, S., Podder, S.: Social Media, Recruitment, Allegiance and the Islamic State. Perspect. Terror. 9, 107–116 (2015).
19. Turk, A.T.: Terrorism and Counterterrorism. In: Goode, E. (ed.) The Handbook of Deviance. pp. 537–548. John Wiley & Sons, Inc, Hoboken, NJ (2015).
20. Jeberson, W., Sharma, L.: Survey on counter Web Terrorism. COMPUSOFT, An Int. J. Adv. Comput. Technol. 4, 1744–1747 (2015).
21. Cheong, M., Lee, V.C.S.: A microblogging-based approach to terrorism informatics: Exploration and chronicling civilian sentiment and response to terrorism events via Twitter. Inf. Syst. Front. 13, 45–59 (2011).
22. Sutton, J., Palen, L., Shklovski, I.: Backchannels on the Front Lines: Emergent Uses of Social Media in the 2007 Southern California Wildfires. In: Proc. ISCRAM, Washington, (2008).
23. Hussain, G., Saltman, E.M.: Jihad Trending: A Comprehensive Analysis of Online Extremism and How to Counter it. Quilliam (2014).
24. Al-Rawi, A.: Anti-ISIS Humor: Cultural Resistance of Radical Ideology. Polit. Relig. Ideol. 7689, 1–17 (2016).
25. Saltman, E.M., Russell, J.: White Paper – The role of prevent in countering online extremism. Quilliam (2014). https://www.quilliamfoundation.org/
26. Wegener, C.: Inhaltsanalyse. In: Mikos, L. and Wegener, C. (eds.) Qualitative Medienforschung: Ein Handbuch. pp. 200–208. UVK, Konstanz (2005).
27. Strauss, A.L.: Qualitative Analysis for Social Scientists. Cambridge Press (1987).
28. Ratkiewicz, J., Conover, M.D., Meiss, M., Gonc, B., Flammini, A., Menczer, F.: Detecting and Tracking Political Abuse in Social Media. In: Proceedings Conference on Weblogs and Social Media. pp. 297–304 (2011).
29. Reuter, C., Ludwig, T., Kotthaus, C., Kaufhold, M.-A., Radziewski, E. von, Pipek, V.: Big Data in a Crisis? Creating Social Media Datasets for Emergency Management Research. i-com J. Interact. Media. 15, (2016).
30. Palen, L., Vieweg, S., Sutton, J., Liu, S.B., Hughes, A.L.: Crisis Informatics: Studying Crisis in a Networked World. In: Proceedings of the International Conference on E-Social Science, Ann Arbor, USA (2007).
31. Stieglitz, S., Dang-Xuan, L., Bruns, A., Neuberger, C.: Social media analytics - An Interdisciplinary Approach and Its Implications for Information Systems. BISE 6 (2014).
32. Rohde, M., Aal, K., Misaki, K., Randall, D., Weibert, A., Wulf, V.: Out of Syria: Mobile Media in Use at the Time of Civil War. Int. J. HCI. 32, 515–531 (2016).
33. Reuter, C., Geilen, G., Gellert, R.: Sicherheit vs. Privatsphäre: Zur Akzeptanz von Überwachung in sozialen Medien im Kontext von Terrorkrisen. In: Mayr, H.C. and Pinzger, M. (eds.) Informatik 2016. LNI, Klagenfurt (2016).

An Interdisciplinary Review of Investor Decision-Making in Crowdfunding

Andreas Hoegen, Dennis M. Steininger, and Daniel Veit

University of Augsburg, Augsburg, Germany
hoegen@is-augsburg.de,
{dennis.steininger,veit}@wiwi.uni-augsburg.de

Keywords: *crowdfunding, decision-making, literature review*

Crowdfunding is a quickly expanding phenomenon that emerged as an alternative way to finance ventures or individuals about a decade ago. In crowdfunding, individuals or ventures use online platforms to collect funds from a relatively large group of investors [1]. Today, it provides increasing competition for traditional financing agents such as venture capitalists (VCs), business angels (BAs), and banks on the one hand [2]. On the other hand, it offers new opportunities to individuals and entrepreneurs in need for financing [3]. The global funding volume was over $34 bn. in 2015 with a growth of more than 1,000% in 3 years according to a recent industry study [4]. Low entry barriers stimulate this growth: in contrast to highly regulated traditional financing markets, investing money through digital channels on crowdfunding platforms is open to almost anyone with an internet connection [3]. But when amateurs become investors, a lack of education and experience in financial decision-making can turn into an expensive endeavor, especially as human decisions frequently seem irrational by economic standards [5, 6].

How investors make decision in traditional startup financing or bank loans is well researched. For crowdfunding, however, prior research covers many individual factors of investment decisions, but does not provide an integrated view on those factors. A better understanding of how investment decisions are made in crowdfunding settings will help crowdfunding concepts and platforms evolve. It can also support investors to make better investment decisions and avoid potentially costly choices. Furthermore, better understanding decision-making in crowdfunding can provide insights into how information systems change the way financial decisions are made. Therefore the research question we address is: *Which factors influence investor decision-making in crowdfunding?*

To answer this question, we conducted a systematic and interdisciplinary literature review on the broad body of crowdfunding studies published to date. For this review we followed Webster & Watson's [7] and Okoli and Schabram's [8] guides for literature reviews. We narrowed our initial selection of 785 crowdfunding articles down to 69 articles that research decision-making in crowdfunding. Using a concept matrix [7] we extracted all factors that do or might influence crowdfunding investment decisions. We inductively developed more abstract clusters by grouping

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Hoegen, A.; Steininger, D.M.; Veit, D. (2017): An Interdisciplinary Review of Investor Decision-Making in Crowdfunding, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 664-667

similar influencing factors. Instead of using pre-defined categories, we allowed the clusters to emerge and develop during our analysis [9]. We continuously discussed and developed the clusters to ensure exhaustive coverage of all factors as well as mutually exclusive clusters with minimal overlap [10]. We used those clusters to build a crowdfunding decision-making framework. During the analysis of our data and concept matrix, six clusters of influencing factors emerged:

Outcome & Quality: the basic value proposition or promised outcome, as well as the (perceived) quality of product, project, and process.

Financial Risk & Campaign Statistics: the financial risk involved in an investment as well as information about the status and characteristics of the crowdfunding campaign.

Founder Perception & Attributes: all characteristics of the founders and how they are perceived by the investors. This includes the basis on which investors derive or judge those qualities.

Social, Community & 3rd Parties: any kind of relations to and behavior of third parties that potentially influence the investment decision.

Context: all attributes of the decision-making context that should not influence a rational investor but impact human decision-makers.

Investor Characteristics: Everything about the investors and their attributes, such as their personal traits and preferences, motives, or involvement.

The framework allows to analyze the many influencing factors from the extant literature and structure our findings. It also enables future discussion and research to build on a more abstract basis, instead of handling a multitude of small factors. We describe the influence of each cluster and the associated factors on decision-making based on findings and example from the analyzed literature.

Based on our analysis and the extant literature, we elicit systematic differences between crowdfunding and traditional investments, such as venture capital or bank loans. The first difference is an even higher impact of social capital that is particularly critical to trigger investment dynamics through herding behavior [12]. Second, crowdfunding investors use many different information sources to substitute a lack of verified data, personal access to the founders and their own experience. These alternative source rank from behavior of others [11, 12], over soft factors (e.g., emotional reactions) [13, 14], to campaign statistics provided by the crowdfunding platform [1]. Finally, the different and more ubiquitous contextual presence of digital technology influences decisions, e.g., through features and functionality of the platform [59], or dependent on the channel investors' use to access a crowdfunding campaign [43].

Literature suggests that the use of information systems as foundation for crowdfunding is the major driver for those systematic differences (e.g., [15–17]). For once, the reach of crowdfunding platforms is high and entry barriers are often low or even nonexistent. Thus, amateurs with little experience regarding such decisions can participate [18, 19]. At the same time, the increased reach and interconnection of digital crowdfunding platforms with other information systems (e.g., social network sites) makes every investment activity highly transparent, thus fostering herding behavior [11]. Second, computer-mediation can increase the perceived (social) distance between founder and investor and increase the propensity to lie [20]. In contrast to professional investors, amateurs have little access to reliable data, collateral, or the chance to meet the founder in person [18, 21]. Hence, they need other means to judge an investment opportunity. Lastly, the interaction of crowdfunding founders and investors is guided by information systems. The platforms create a digital context around the transaction that determines how both parties interact with each other. In addition, digital interfaces guide each investor's decision process [22]. Therefore this context influences the decision.

Our findings lead to different takeaways for IS researchers and practitioners. A better understanding of the IS specific factors in crowdfunding decision-making would be beneficial, e.g., how interfaces and platform functionality influence decisions on crowdfunding platforms. On this basis, platform design could be re-evaluated and interfaces improved to help investors decide in their own best interest. In addition, focusing on the entire decision-making process and how different influencing factors and clusters interact could lead to insights on why and how investors decide. Many extant studies focus on assessing correlations between single factors and the funding success of a campaign. Investigating the whole decision-making process in detail and how information systems influence it could build on those findings and lead to additional insights. Experimental and design studies could be useful approaches to elicit better ways to build platforms and learn how to create interactions with information systems in investment situations. Lastly, more research into traditional investments versus crowdfunding could help to narrow down which deviations in investment behavior are due to differences in the concepts (one or few professional investors versus large crowd) and which are rooted in the usage of information technology to implement the concept.

Crowdfunding offers new ways of financing ventures as well as new opportunities to support others and invest money. Given the historic growth rate and the size of the market for loans and investments, its impact on credit markets, venture capital, and entrepreneurial dynamics will likely further expand. We believe that a better understanding of how decisions are made in this context will benefit all involved parties and help mitigate bad choices for investors and founders. The authors hope that their results can bring inspiration to future research and thereby support the development of the field.

References

1. Mollick, E.: The dynamics of crowdfunding: An exploratory study. Journal of Business Venturing. 29, 1–16 (2014).
2. Khavul, S., Chavez, H., Bruton, G.D.: When institutional change outruns the change agent: The contested terrain of entrepreneurial microfinance for those in poverty. Journal of Business Venturing. 28, 30–50 (2013).
3. Bruton, G., Khavul, S., Siegel, D., Wright, M.: New Financial Alternatives in Seeding Entrepreneurship: Microfinance, Crowdfunding, and Peer-to-Peer Innovations. Entrepreneurship Theory and Practice. 39, 9–26 (2015).

4. Massolution: 2015CF - The Crowdfunding Industry Report, http://www.crowdsourcing.org/editorial/global-crowdfunding-market-to-reach-344b-in-2015-predicts-massolutions-2015cf-industry-report/45376, (2015).
5. Kahneman, D.: Thinking, fast and slow. Macmillan (2011).
6. Thaler, R.H.: From Homo Economicus to Homo Sapiens. The Journal of Economic Perspectives. 14, 133–141 (2000).
7. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly. 26, xiii–xxiii (2002).
8. Okoli, C., Schabram, K.: A Guide to Conducting a Systematic Literature Review of Information Systems Research. 10, (2010).
9. Glaser, B., Strauss, A.: The discovery of grounded theory. London: Weidenfeld and Nicholson. 24, 288–304 (1967).
10. Wolfswinkel, J.F., Furtmueller, E., Wilderom, C.P.: Using grounded theory as a method for rigorously reviewing literature. European Journal of Information Systems. 22, 45–55 (2013).
11. Burtch, G., Ghose, A., Wattal, S.: An Empirical Examination of the Antecedents and Consequences of Contribution Patterns in Crowd-Funded Markets. Information Systems Research. 24, 499–519 (2013).
12. Colombo, M.G., Franzoni, C., Rossi-Lamastra, C.: Internal social capital and the attraction of early contributions in crowdfunding. Entrepreneurship Theory and Practice. 39, 75–100 (2015).
13. Dorfleitner, G., Priberny, C., Schuster, S., Stoiber, J., Weber, M., de Castro, I., Kammler, J.: Description-text related soft information in peer-to-peer lending – Evidence from two leading European platforms. Journal of Banking & Finance. 64, 169–187 (2016).
14. Genevsky, A., Knutson, B.: Neural Affective Mechanisms Predict Market-Level Microlending. Psychological Science (Sage Publications Inc.). 26, 1411–1422 (2015).
15. Liu, D., Brass, D.J., Lu, Y., Chen, D.: Friendships in Online Peer-to-Peer Lending: Pipes, Prisms, and Relational Herding. MIS Quarterly. 39, 729–742 (2015).
16. Moritz, A., Block, J., Lutz, E.: Investor communication in equity-based crowdfunding: a qualitative-empirical study. Qualitative Research in Financial Markets. 7, 309–342 (2015).
17. Ordanini, A., Miceli, L., Pizzetti, M., Parasuraman, A.: Crowd-funding: transforming customers into investors through innovative service platforms. Journal of Service Management. 22, 443–470 (2011).
18. Ahlers, G.K.C., Cumming, D., Günther, C., Schweizer, D.: Signaling in Equity Crowdfunding. Entrepreneurship Theory and Practice. 39, 955–980 (2015).
19. Ley, A., Weaven, S.: Exploring Agency Dynamics of Crowdfunding in Start-up Capital Financing. Academy of Entrepreneurship Journal. 17, 85–110 (2011).
20. Whitty, M.T., Joinson, A.: Truth, lies and trust on the Internet. Routledge (2008).
21. Chen, X.-P., Yao, X., Kotha, S.: Entrepreneur Passion and Preparedness in Business Plan Presentations: A Persuasion Analysis of Venture Capitalists' Funding Decisions. Academy of Management Journal. 52, 199–214 (2009).
22. Thaler, R.H., Sunstein, C.R.: Nudge: Improving Decisions About Health, Wealth and Happiness. Yale University Press, New Haven (2008).

Helper, Sharer or Seeker? – A Concept to Determine Knowledge Worker Roles in Enterprise Social Networks

Janine Hacker[1], Freimut Bodendorf[1], and Pascal Lorenz[2]

[1] Friedrich-Alexander-Universität Erlangen-Nürnberg, Institute of Information Systems, Nürnberg, Germany
{janine.hacker,freimut.bodendorf}@fau.de
[2] University of Haute Alsace, IUT, Colmar, France
{pascal.lorenz}@uha.fr

Abstract. In order to manage knowledge work, companies need to understand how knowledge is shared, integrated, translated and transformed in organisational practice. However, knowledge work often happens in informal organisational structures, thus, making it difficult to identify and understand the occurring knowledge practices and participating actors. Enterprise Social Networks (ESN), i.e. internally accessible social networking services, have evolved as important platforms for knowledge work. Facilitating knowledge interactions between users, the analysis of ESN data might be well suited for characterising and identifying knowledge actions and different knowledge worker roles. Drawing on an existing knowledge worker role typology as well as findings from social media research, this paper develops a conceptual basis that serves as starting point for determining knowledge worker roles using ESN data. The next steps of this research involve the empirical testing of the typology using data obtained from a real case scenario.

Keywords: Enterprise Social Networks, knowledge worker, social roles, knowledge work.

1 Introduction

While knowledge has been recognised as a key source of competitive advantage, the management of knowledge work is a challenging – and often unsuccessful – task [1]. One reason for knowledge management (KM) initiatives to fail lies in the very nature of knowledge and knowledge work. As such, knowledge work is often conducted in informal organisational structures that exist next to the formal hierarchy and are not fully obvious to the company's management nor to the involved actors [2]. Being emergent and contextual, knowledge and knowledge work are difficult to describe and to measure, and therefore, difficult to manage.
To be able to (better) manage knowledge work, organisations need to understand the components of knowledge work, i.e. knowledge actions, and the different knowledge worker roles performing these actions [1]. Research dealing with informal networks and key roles within these networks [e.g. 2] as well as existing knowledge worker role

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Hacker, J.; Bodendorf, F.; Lorenz, P. (2017): Helper, Sharer or Seeker? - A Concept to Determine Knowledge Worker Roles in Enterprise Social Networks, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 668-682

typologies [e.g. 3] provide insights regarding different actor roles shaping knowledge work in organisations. Yet, these studies determine roles based on survey data. Thus, the significant manual effort involved in the data collection and analysis makes it difficult to recognise knowledge worker roles in larger populations.

More and more organisations use Enterprise Social Networks (ESN) to support collaboration and knowledge sharing. The users of these platforms engage in use practices such as discussing, asking questions, giving advice, and sharing content [4]. Doing so, they leave a number of digital traces stored in the back end of the ESN [5]. Reflecting knowledge (inter)actions of ESN users, the record of user activity may be well suited for identifying different knowledge worker roles in organisations.

Drawing on the knowledge worker role typology by [3], this paper establishes a conceptual basis for determining knowledge worker roles using ESN data. Based on a systematic literature review, ESN use cases are analysed and compared with the knowledge actions specified in [3]. The knowledge actions are operationalised using a set of metrics characterising user activity, i.e. quantitative and qualitative information on the users' posting behaviour, in ESN. As a result of the literature analysis, the paper proposes a revised typology which will be empirically tested in a follow-up project.

This paper contributes to KM research by presenting the initial steps for determining knowledge worker roles in ESN. It advances the field of ESN data analytics by proposing a set of metrics to characterise ESN user behaviour. As for practitioners, awareness of the knowledge worker roles that employees take on can improve decision-making at the intersection of human resources management (HRM) and KM.

2 Theoretical Background and Related Work

This paper draws on KM research and social media research to derive a concept to determine knowledge worker roles in ESN. The following sections contrast perspectives on knowledge and introduce research on KM-related roles in organisations. Section 2.3 deals with ESN features and ESN data. Also, selected works on the identification of roles in online social spaces are presented.

2.1 Knowledge and Knowledge Work

Two main perspectives regarding the study of knowledge can be distinguished: The possession perspective and the practice perspective [6]. Scholars taking up the possession perspective consider knowledge as an object that is held by an individual knower [e.g. 7]. Accordingly, KM strategies focus on the extraction and explication of individual knowledge in order to make it accessible to others [8]. On the other hand, researchers adopting the practice perspective argue knowledge to be created and shared through social interactions between individuals [e.g. 9]. Consequently, this perspective suggests the management of knowledge work, i.e. activities "characterised by an emphasis on theoretical knowledge, creativity and use of analytical and social skills" [10, p. 773], rather than managing knowledge [1]. In particular, knowledge work should be stimulated and enhanced, e.g. by connecting individuals from different departments to reinforce knowledge sharing. This paper adopts the knowledge-in-practice perspective since it more adequately reflects how knowledge work occurs on ESN. As such, ESN offer a conversational space for knowledge work rather than a place for storing knowledge [4]. Also, the knowledge-in-practice perspective reflects the dynamics and complexity of contemporary work settings [e.g. 1].

2.2 Knowledge Worker Roles in Organisations

Individuals performing knowledge work, so-called knowledge workers, are primarily occupied with 'non-routine' problem solving [3]. While knowledge work is an important topic in the KM literature, little is known about the different roles that knowledge workers assume during their daily work [3]. In this regard, some role typologies [e.g. 11] link roles to specific KM-related processes, such as *finding*, *packaging*, and *distributing* knowledge. The corresponding roles are named *knowledge packagers* and *knowledge distributors*, for instance [11, p. 28-30]. Yet, linking only one knowledge action with a certain role, these typologies do not seem to adequately reflect the complexity of knowledge work and implicitly treat knowledge like an object. Thus, they are not directly applicable for determining knowledge worker roles in ESN.
Reflecting the knowledge-in-practice perspective, studies investigating informal organisational structures [e.g. 2] offer insights regarding actor roles involved in knowledge work, among them *central connectors*, *boundary spanners* and *peripheral specialists* [e.g. 12]. Moreover, [3] propose a typology including ten knowledge worker roles (cf. [3] for a detailed description of the roles), each of which is associated to different degrees with a set of knowledge actions out of 13 specified knowledge actions (Figure 1).

Knowledge action / Knowledge worker role	Acquisition	Analyse	Authoring	Co-authoring	Dissemination	Expert Search	Feedback	Information organisation	Information search	Learning	Monitoring	Networking	Service Search
Controller	o	++	+	-	+	-	+	o	o	--	o	--	--
Helper	o	++	-	-	++	o	++	o	++	o	-	o	--
Learner	+	++	-	-	--	o	-	+	++	+	o	o	-
Linker	+	++	o	o	o	-	-	o	++	o	o	o	--
Networker	o	o	-	-	+	+	o	o	o	o	o	++	--
Organizer	o	++	-	--	o	-	o	+	o	-	+	+	o
Retriever	++	+	-	--	-	o	-	+	+	o	+	o	o
Sharer	-	o	+	+	++	--	o	o	o	-	-	+	-
Solver	o	++	-	--	o	o	o	o	+	o	-	-	-
Tracker	o	+	-	--	o	o	o	+	+	o	++	o	o

++ Very important + Important o Undecided - Less important -- Not important

Figure 1. Knowledge worker roles and associated knowledge actions (based on [3])

In this regard, Figure 1 depicts the relationships between knowledge worker roles and knowledge actions and indicates the importance of a knowledge action for a certain role [3]. In line with the knowledge-in-practice perspective adopted in this paper, the knowledge worker role typology by [3] reflects roles identified in informal organisational networks, e.g. *boundary spanners*. Derived based on a literature review as well as an empirical study, the typology includes a comprehensive set of roles, associates specific knowledge actions with each role and indicates the importance of an action for a role. As ESN have evolved as platforms for knowledge work [5], users may perform knowledge actions similar to those specified by [3] and take up similar roles. Thus, the knowledge worker role typology by [3] serves as the basis for determining roles in ESN.

2.3 Functionalities and Analysis of Enterprise Social Networks

In recent years, companies have started to use internal online social networks, i.e. ESN, well-known examples of which include IBM Connections, Jive or Yammer [13]. Relying on Web 2.0 technology, ESN are web-based Intranet platforms that generally offer the following features [19-20]:

- *Profile pages* allowing users to present information about themselves
- *Following* other users to see their updates
- *Activity streams* displaying updates from other users and followed topics within an integrated newsfeed
- *Searching* the content stored on the ESN
- *Group capabilities* that allow interactions within public or restricted groups
- *Discussion threads* where users can start conversations via status updates, share files and participate in conversations by replying to, liking, rating and sharing the messages of other users
- *Tagging* of other users or topics in messages
- *Bookmarks* that allow for saving, organising and sharing content
- *Blog and wiki capabilities* that enable the (collaborative) creation of content and storing of information
- *Social analytics* that provide users with contact recommendations

The actions of users on ESN are visible to other users and persist over time [16]. Specifically, communicative actions are stored as digital traces, i.e. "digitally stored, event-based, chronological records of activities of actors" [5, p. 4], in the ESN back end. The accumulated data can generally be classified according to the following categories [17]: *Activities* (usage data), *content* (user-generated data), and *relations* (structural data). A combination of quantitative and qualitative methods facilitates the analysis of ESN data [17]. Usage data enables the development of metrics, e.g. *number of status updates created (per month)*, that quantify user activities. The content of communications, e.g. the content of a status update, can be analysed using qualitative content analysis, for instance to identify communication genres [18], or (partly) automatically, for instance using text mining techniques. Also, different kinds of relations can be inferred from user interactions on the platform, e.g. based on *following* relationships [17]. Structural data can be analysed using social network analysis (SNA) metrics that enable characterising the position of individual actors, e.g. based on centrality measures [19].

2.4 Roles in Online Social Spaces

Interactions between individuals in online social spaces, such as discussion boards, result in communication patterns. Roles can then be described by “structural signatures” reflecting their position in the social network and behavioural patterns characterising their participation behaviour, e.g. how often they contribute to a discussion [20, 21].
Compared to the high number of articles dealing with user roles in public online settings [22], research on user roles in ESN is still sparse. A case study by [23] identifies *discourse drivers*, who primarily disseminate topics, and *information retrievers*, who are interested in finding and using information, as two mutually interdependent actor roles. Other studies analysing knowledge exchange differentiate between *knowledge seekers* and *knowledge contributors* [24], and between *givers*, *takers*, and *matchers* respectively [25]. Additionally, *value adding users*, i.e. users who contribute and share their knowledge in the ESN, are suggested as ESN actor role [26]. Analysing log data, [27] differentiate between *active contributors*, *moderate contributors*, and *readers* as well as *active* and *occasional* users. All in all, the role concepts suggested in ESN research are not very detailed since they are based on relatively few behavioural dimensions and identified using a limited set of metrics. However, metrics used in studies investigating roles in public online settings can inform the development of metrics to determine the knowledge actions suggested by [3] in ESN.

3 Determining Knowledge Worker Roles in Enterprise Social Networks

This paper uses the knowledge worker role typology by [3] as a starting point for determining knowledge worker roles in ESN. While the original typology is based on a literature analysis and a survey, this paper addresses the question how the typology needs to be adapted in order to facilitate the determination of knowledge worker roles using ESN data.
The adaptation of the typology involves three steps: The first step (section 3.1) concerns the identification of ESN use cases that match the knowledge actions specified in [3]. The goal is to find out whether the knowledge actions are reflected in ESN use cases as well as to assess if they can be determined as distinct actions using ESN data. The second step (section 3.2) focuses on the operationalisation of the knowledge actions by developing a set of ESN metrics. Following the approach of explicit role discovery [22], the knowledge actions associated with a role serve as criteria for individual users to be assigned to a certain role as described in section 3.3.

3.1 Applicability of Knowledge Actions in Enterprise Social Networks

Focusing on information systems research on ESN, the literature review performed in step 1 considers studies published in major information systems journals and conferences according to the Association for Information Systems [28] and the rating of the German Academic Association for Business Research [29]. Covering the years 2005-

2015, multiple search terms and combinations of them were considered, among them microblogging, enterprise social network, enterprise social media, social software, and KM. A detailed description of the literature review process and assignment of the selected papers to six metatopics is described in [30]. The contents of the following section are compiled based on the publications primarily assigned to the metatopic "ESN usage and behaviour" as well as on a follow-up search for papers dealing with ESN usage covering the year 2016. Specifically, papers assigned to this metatopic analyse the ways and purposes of using ESN in organisations, thus identifying ESN use cases. Analysing the selected 28 publications in detail, ESN use cases mentioned in the papers, such as *input generation* or *information storage* [18], are matched with the 13 knowledge actions according to their descriptions in [3]. The ESN use case *information storage,* for instance, is matched with the knowledge action *information organisation* since it is related to managing information on the ESN for future reference. Table 1 consolidates the findings of the literature analysis. The subsequent paragraphs detail which and to what extent the different knowledge actions are applicable to ESN.

Acquisition describes the collecting of information with a conscious goal, e.g. to develop skills or advance a project [3]. Actions related to information gathering on ESN have been identified in a number of studies [e.g. 18]. For instance, acquisition of knowledge is explicitly referred to in the definition of "consumptive" ESN use in [31]. However, the second part of the definition of *acquisition* [3], i.e. having a conscious goal, is not explicitly addressed in the existing ESN literature and appears difficult to determine based on ESN data.

Analyse is defined as carefully examining or thinking about something with the objective of understanding it [3]. The action *analyse* is not necessarily linked with a piece of written information that could be submitted and thus, become visible, on an ESN. While the knowledge action *analyse* is not explicitly mentioned in the ESN literature, the analysis of content can be assumed to be part of the knowledge actions *learning* as well as *monitoring*. Thus, users are concluded to analyse content on the ESN by reading it in order to learn or keep themselves up-to-date about something.

Authoring refers to the creation of information objects (by one person), e.g. textual content, using a software system [3]. As such, *authoring* applies to all textual content submitted to the platform, including posts to the main stream. To better differentiate it from other knowledge actions concerning the ESN main stream (e.g. *dissemination* and *feedback*), in this paper, *authoring* is conceived of as initiating conversations, e.g. to propose an idea [e.g. 4], rather than contributing to existing conversations. Moreover, *authoring* is related to the creation of wiki entries, notes or blog posts [32, 33] which are features of most ESN platforms.

Contrary to *authoring*, **co-authoring** means the collaborative creation of content [3]. *Co-authoring* occurs on ESN platforms when users edit the content created by other users, e.g. when updating content to a new version [34], editing a wiki page [32], articles [25], or blog posts [33].

Dissemination refers to the spreading of information or information objects, such as work results [3]. With regard to ESN, employees share factual information in status updates, e.g. information about objects or people [35], and event notifications [18, 36]. Secondly, users employ ESN to share information objects to make them accessible to others [34], e.g. files in general, project descriptions [37] as well as links to content that already exists [18, 34]. Thirdly, individuals post updates [e.g. 4], regarding their work activity, mood, or physical location [38] as well as "me-statements" that reveal something about them [35]. As a means for impression management [39], users share personal information and information about their professional background.

Expert search refers to retrieving an expert in order to discuss and solve a specific problem [3]. In ESN, retrieving experts is considered as one form of knowledge seeking, i.e. seeking knowledge about people with specific characteristics [24], such as particular skills or connections [4, 18, 40].

The knowledge action **feedback** is described as the evaluation of a proposal or an information object [3]. As for ESN, feedback is part of conversations, e.g. discussions of different options for action to solve a problem [e.g. 18]. Additionally, "social feedback" is recognised as a form of informal communication on ESN [34]. In terms of ESN features, *feedback* can be expressed in liking, rating [34] or commenting on the content created by another user [32] as well as by voting the posted content up or down [41].

The personal or organisational management of collected information is referred to as **information organisation** [3]. In this regard, ESN are used to store information [18, 36]. Furthermore, the labelling of content with a topic [34], e.g. using a tag, can be concluded as a form of *information organisation*.

Information search is described as looking up information on a particular topic and in a specific form [3]. ESN users can "look up" information stored on the platform using the search feature [34]. Moreover, status messages containing questions and information requests are a dominant use case in the ESN literature [e.g. 31] and can be considered as an "explicated" form of *information search*. For instance, employees seek factual knowledge, opinions, and recommendations in status updates [24].

Learning refers to acquiring new knowledge, skills or understanding while performing work tasks, in exchange with others, or using formalised learning material [3]. As such, *learning* appears to be closely related to the action *acquisition* and could be assumed to be one motive of knowledge actions related to search. *Learning* on ESN occurs via a mix of consuming and contributing activities. Browsing the record of activities in project groups, for instance, enables new team members to learn about the history and current status of a project [42]. Also, ESN enable visibility into other users' conversations [16]. As users read and engage in these conversations, they learn about people's expertise and what is important in the organisation [4]. Therefore, ESN facilitate continuous (social) organisational learning [16, 18].

Monitoring is described as keeping oneself informed about selected topics, e.g. using different electronic information resources [3]. Just like the actions *analyse* and *learning*, *monitoring* is difficult to observe based on what employees explicitly post to the network. Yet, ESN enable *monitoring* by supporting different kinds of situation awareness, that is knowing who is doing what (activity awareness), knowing relevant contact persons (structure awareness), as well as knowing what other people are interested in and knowing who is talking to whom (social awareness) [43]. In particular, *monitoring* is facilitated by features related to subscribing to users' updates or topics [19-20]. *Monitoring* may be closely related to *learning* since learning can be an outcome of monitoring activities.

Table 1. Overview of knowledge actions and associated ESN use cases

Knowledge action	**Associated ESN use case and sample references**	**Concerned ESN feature** [14, 15]
1 Authoring	*1a* Initiating a conversation to ask for input / ideas [18]; *1b* Creating a wiki page / blog post [32]	Status updates; Wiki / Blog
2 Co-authoring	*2* Editing of a wiki page / blog post [32, 33]	Wiki / Blog
3 Dissemination	*3a* Providing updates (e.g. on work activity) [4, 39]; *3b* Sharing of files / references [36]	Status updates; File sharing
4 Expert search	*4a* Retrieving an expert using search feature [34]; *4b* Requesting for an expert within a status update [44]	Search; Status updates (expert seeking)
5 Feedback	*5a* Providing feedback within conversations [18]; *5b* Providing social feedback [36]; *5c* Liking or rating of content [34]	Discussion thread (participation); Rating / liking
6 Information organisation	*6a* Saving of meeting minutes [18]; *6b* Bookmarking content [26]; *6c* Labelling / tagging of content [34]	File repository; Bookmarks; Tagging
7 Information search	*7a* Retrieving information using search feature [18, 34]; *7b* Asking for information in a status update [31, 36]	Search; Status updates (information seeking)
8 Learning	*8a* Reading past conversations [16]; *8b* Participating in a conversation [4]	Activity stream (browsing); Discussion thread (participation)
9 Monitoring	*9a* Reading past conversations [43]; *9b* Subscribing to users' messages or keywords [36]	Activity stream (browsing); Following other users / topics
10 Networking	*10a* Following other users [32]; *10b* Creating social relations [37]; *10c* Tagging other users [32]; *10d* Talking about non-work related matters [4]; *10e* Commenting on status updates [37]	Following other users; Discussion thread (participation); Discussion thread (social use)
11 Service search	*11* Asking for a solution in a status update [45]	Status updates (solution seeking)

Interactions with other people and organisations in order to exchange information and to establish contacts are described as **networking** [3]. While every interaction between users includes some form of information exchange, *networking* is concluded as "social use" of ESN [31] in this study. As such, it describes communication aimed at the maintenance of existing and creation of new social relationships in order to build personal context [44]. Messages indicating social use are often related to informal communications [e.g. 18]. For instance, the tagging of other users in messages [32] or the use of emoticons [35] could serve as an indicator for *networking*.

Service search describes the seeking for specialised web services that offer particular functions in order to address a given problem. Compared to *information* and *expert search*, the object to be retrieved is different [3]. On ESN, messages related to seeking solutions or specific resources could be concluded as *service search* [45].

Summarising the findings of the literature analysis, Table 1 lists ESN use cases (citing sample references) associated with the knowledge actions as well as ESN features employed to engage in the use cases. In this regard, one knowledge action can be described by several ESN use cases. In contrast to the knowledge actions by [3], Table 1 does not include the actions *analyse* and *acquisition*. In this regard, no use case could be identified for *analyse* since this action is not explicitly mentioned in the ESN literature. Yet, it is implicitly included in the actions *learning* and *monitoring* as in reading content in order to learn or keep informed respectively. Moreover, *acquisition* is merged with *information search* into the action *information search*. On the one hand, only the component of *acquisition* related to *information search* can be observed on ESN. On the other hand, users are assumed to generally have a goal when searching for information.

3.2 Determining Knowledge Actions in Enterprise Social Networks

Having identified the ESN use cases associated with a particular knowledge action, metrics for recognising knowledge actions using ESN data are designed. As such, the metrics shown in Table 2 reflect the concerned ESN features (Table 1) and facilitate the quantification of the ESN use cases connected to a knowledge action. The numbering of the metrics corresponds to the numbering of the ESN use cases in Table 1. The metrics are designed based on existing studies [e.g. 32, 34] and in accordance with the data generally available in ESN [17]. Moreover, Table 2 indicates the categories of ESN data (section 2.3), i.e. activities (usage data), content (user-generated data), and relations (structural data) [17], required to implement the metrics. The knowledge action *dissemination*, for instance, is connected to the use cases *providing updates* and *sharing of files* in ESN (Table 1). Exemplary metrics to determine the extent to which a user engages in the knowledge action *dissemination* then include the *number of status updates* as well as the *number of status updates that contain an attachment*. These metrics require the collection and analysis of data from the category *activities*, i.e. usage data. Furthermore, Table 2 indicates the complexity and hence effort required to analyse the different knowledge actions using ESN data. In this regard, knowledge actions in rows coloured in grey are recognised as less difficult to analyse than the ones in rows without shading. The effort involved in analysing a knowledge action depends on the data dimensions and data analysis methods (cf. section 2.3) required to implement the respective metrics.

Table 2. Operationalisation and analysis of knowledge actions

Knowledge action [3]	**Example metrics to quantify ESN use cases** (cf. Table 1)	**Data category** [17]
1 Authoring	*1a* No. of status updates (first messages); *1b* No. of wiki / blog entries created	Activities
2 Co-authoring	*2* No. of wiki / blog entries edited	Activities
3 Dissemination	*3a* No. of status updates (first messages); *3b* No. of status updates (with links / files)	Activities
4 Expert search	*4a* No. of searches performed (person's name); *4b* No. of questions asked (expert seeking)	Activities; Content
5 Feedback	*5a* No. of replies in threads containing discussions / feedback; *5b* No. of praise messages; *5c* No. of ratings performed / likes given	Activities; Content
6 Information organisation	*6a* No. of files saved; *6b* No. of documents bookmarked *6c* No. of tags created; *6d* No. of documents tagged;	Activities
7 Information search	*7a* No. of searches performed; *7b* No. of questions asked (information seeking)	Activities; Content
8 Learning	*8a* Time spent browsing content; *8a* No. of conversations clicked on; *8b* No. of threads participated in	Activities
9 Monitoring	*9a* Time spent browsing content; *9b* No. of users followed; *9b* No. of topics followed	Activities
10 Networking	*10a* No. of users followed; *10b* No. of new following relationships in a certain period of time; *10c* Out-degree (tagging relationships); *10d* No. of messages (social use) *10e* Out-degree (reply relationships);	Activities; Content; Relations
11 Service search	*11* No. of questions asked (solution seeking)	Activities; Content

Thus, metrics indicating *authoring*, *co-authoring*, *dissemination*, *information organisation*, *learning*, and *monitoring* can be implemented based on data of the category *activities*. Having obtained this data, the corresponding metrics can be calculated in a database or using a spreadsheet software straightaway. Contrary, implementing metrics to describe the knowledge actions in the white rows requires a mix of qualitative and quantitative methods [17]. For instance, metrics indicating actions related to search, such as *expert search* or *information search*, are based on analysing search queries of users (using the search feature) as well as based on questions in status updates. Hence, the message content needs to be considered to decide *whether* a status update is a question and *what* the user is asking for. Scanning of messages for question marks and question words, e.g. “who” or “how” [46], or applying text mining techniques could help partly automate this task. Having identified questions and their focuses, it is possible to determine *how often* a user has created a question. A mixed methods approach is also suggested for *feedback* which can be recognised by analysing the content of messages regarding the provision of (social) feedback and / or determining how often a user has liked or rated another user’s content. Finally, metrics indicating *networking* need to be calculated using log file analysis and SNA. SNA is usually performed using specialised tools, e.g. UCINET, that require the data to be prepared in a certain way.

3.3 Determining Knowledge Worker Roles in Enterprise Social Networks

Besides including conceptually overlapping knowledge actions (section 3.1), the knowledge worker role typology by [3] contains several roles that are associated with the same knowledge actions to similar extents. In particular, this applies to the roles *learner* and *retriever* as well as *organiser* and *tracker* (Figure 1). Engaging in the same knowledge actions to similar degrees, these knowledge worker roles appear difficult to distinguish based on a mainly quantitative analysis of ESN data. Therefore, *learner* and *retriever* are merged into a new role called *seeker.* Further, *organiser* and *tracker* are consolidated into a role called *coordinator*. The extent to which the new roles are associated with the knowledge actions is determined by comparing and averaging their individual degrees for the different actions as specified in [3]. Figure 2 depicts the revised knowledge worker role typology. In this regard, knowledge actions rated as (very) important or less important / unimportant are recognised as significant for characterising the different knowledge worker roles and should be considered first in order to detect the roles in ESN. Knowledge actions marked with “o” are considered less discriminative for the different roles. For instance, to identify *controllers*, metrics (Table 2) suggested to quantify the respective discriminative knowledge actions (marked with ++ / + and -- / - in Figure 2) should be calculated for all users in the sample. Combining the

Knowledge action / Knowledge worker role	Authoring	Co-authoring	Dissemination	Expert Search	Feedback	Information organisation	Information search	Learning	Monitoring	Networking	Service Search
Controller	+	-	+	-	+	o	o	--	o	--	--
Helper	-	-	++	o	++	o	++	o	-	o	--
Seeker	-	--	--	o	-	+	++	+	o	o	-
Linker	o	o	o	-	-	o	++	o	o	o	--
Networker	-	-	+	+	o	o	o	o	o	++	--
Coordinator	-	--	o	o	o	+	+	-	++	o	o
Sharer	+	+	++	--	o	o	o	-	-	+	-
Solver	-	--	o	o	o	o	+	o	-	-	-

++ Very important   + Important   o Undecided   - Less important   -- Not important

Figure 2. Adapted knowledge worker role typology (based on [3])

information presented in Table 2 and Figure 2, individuals meeting the criteria for the role *controller* would exhibit above average values for metrics associated with the corresponding (very) important knowledge actions (++ / +) and below average values for the ones rated as less important / unimportant (-- / -) as compared to other users in the sample.

4 Discussion and Future Work

Drawing on findings from social media and ESN research, this paper adapts the knowledge worker role typology by [3] to an ESN context. In comparison to the typology by [3], the adapted typology includes a reduced set of knowledge actions and knowledge worker roles. Offering metrics to quantify the knowledge actions, the paper provides a conceptual basis for determining knowledge worker roles using ESN data.

The results of this paper need to be weighted up against its limitations. Firstly, the matching of the ESN use cases with the knowledge actions in [3] required some judgement is thus, not entirely objective. Specifically, there is little information on some of the knowledge actions, e.g. *learning*, in the ESN literature. Hence, some assumptions have to be made how *learning* occurs on ESN and how it can be measured. Secondly, knowledge actions that overlap conceptually and correlate for different roles as well as roles associated with the same knowledge actions to similar extents were merged to avoid redundancy. While this is necessary to enable the determination of (distinct) knowledge worker roles based on a quantitative analysis of ESN data, the adapted typology is less detailed and could miss actions and roles involved in knowledge work.

Following-up on these issues, the next step of this research project concerns the testing of the knowledge worker role typology using ESN data obtained from a real case scenario. The planned empirical analysis has implications for the developed metrics, the suggested roles, as well as for the general feasibility of the approach. First of all, the empirical analysis will show whether the proposed metrics can be computed using ESN data. In this regard, metrics may have to be adapted to suit the specifics of the data set or removed if particular data, for instance, the content of messages, is not available. Also, correlations between metrics assigned to one knowledge action should be empirically tested. For determining the roles, adequate thresholds for the metrics in order to distinguish less / unimportant from (very) important knowledge actions have to be specified. The actual role analysis will show whether the suggested roles are valid and point out if and how the typology needs to be further adapted and refined. On the one hand, overlapping roles could be revealed in case many users meet the criteria of several roles. If, on the other hand, many users cannot be assigned to a role, the current typology would be indicated to miss certain, possibly ESN-specific, roles. In this regard, roles identified in public online social spaces, such as *experts* (who, for instance, receive many questions) or *discussion persons* (who join many conversations) [20] could be considered to extend the typology. In conclusion, the planned empirical testing will indicate to what extent ESN data facilitates determining the specified knowledge worker roles.

5 Conclusion

The management of knowledge work requires an understanding of the performed knowledge actions and involved knowledge worker roles. The identification and measurement of these actions as well as the determination of knowledge worker roles is an important and current challenge in KM research and practice.

This paper contributes to KM research by offering a conceptual basis for determining knowledge actions and knowledge worker roles based on ESN data. The study moreover contributes to the emerging body of ESN data analytics by developing metrics that characterise ESN user behaviour. For practitioners, the detection of different roles can improve organisational knowledge transparency and lead to more evidence-based decisions at the intersection between KM and HRM. Facilitating new opportunities for information sharing and exchange, ESN have the potential to significantly change the way knowledge work is conducted in organisations. It remains to be investigated whether similar knowledge worker roles as suggested by [3] can be recognised using ESN data or if interactions on ESN lead to new roles in the context of knowledge work.

References

1. Newell, S., Robertson, M., Scarbrough, H., Swan, J.: Managing Knowledge Work and Innovation. Palgrave Macmillan, Basingstoke (2009)
2. Allen, J., James, A.D., Gamlen, P.: Formal versus informal knowledge networks in R&D: A case study using social network analysis. R&D Manag. 37, 179–196 (2007)
3. Reinhardt, W., Schmidt, B., Sloep, P., Drachsler, H.: Knowledge Worker Roles and Actions-Results of Two Empirical Studies. Knowl. Process Manag. 18, 150–174 (2011)
4. Riemer, K., Scifleet, P.: Enterprise social networking in knowledge-intensive work practices: a case study in a professional service firm. In: 23rd Australasian Conference on Information Systems, Geelong (2012)
5. Behrendt, S., Richter, A., Riemer, K.: Conceptualisation of Digital Traces for the Identification of Informal Networks in Enterprise Social Networks. In: 25th Australasian Conference on Information Systems, Auckland (2014)
6. Cook, S.D.N., Brown, J.S.: Bridging Epistemologies: The Generative Dance Between Organizational Knowledge and Organizational Knowing. Organ. Sci. 10, 381–400 (1999)
7. Nonaka, I.: A Dynamic Theory of Organizational Knowledge Creation. Organ. Sci. 5, 14–37 (1994)
8. Hansen, M.T., Nohria, N., Tierney, T.: What's your strategy for managing knowledge? Harv. Bus. Rev. 77, 106–116 (1999)
9. Orlikowski, W.J.: Knowing in Practice: Enacting a Collective Capability in Distributed Organizing. Organ. Sci. 13, 249–273 (2002)
10. Frenkel, S., Korczynski, M., Donoghue, L., Shire, K.: Re-Constituting Work: Trends towards Knowledge Work and Info-Normative Control. Work. Employ. Soc. 9, 773–796 (1995)
11. Davenport, T.H.: Thinking for a Living: How to Get Better Performances and Results from Knowledge Workers. Harvard Business Review Press (2005)
12. Parise, S., Cross, R., Davenport, T.H.: Strategies for Preventing a Knowledge-Loss Crisis. MIT Sloan Manag. Rev. 47, 31–38 (2006)

13. Gartner: Gartner Says 80 Percent of Social Business Efforts Will Not Achieve Intended Benefits Through 2015, http://www.gartner.com/newsroom/id/2319215 (Accessed: 25.08.2016)
14. Drakos, N., Mann, J., Gotta, M.: Magic quadrant for social software in the workplace, http://www.gartner.com/technology/reprints.do?id=1-20TBOV4&ct=140903&st=sb (Accessed 31.08.2015)
15. Koplowitz, R.: The Forrester Wave: Enterprise Social Platforms, Q2 2014. http://public.dhe.ibm.com/common/ssi/ecm/lo/en/lol14021usen/LOL14021USEN.pdf (Accessed 27.04.2016)
16. Leonardi, P.M., Huysman, M., Steinfield, C.W.: Enterprise Social Media: Definition, History, and Prospects for the Study of Social Technologies in Organizations. J. Comput. Commun. 19, 1–19 (2013)
17. Behrendt, S., Richter, A., Trier, M.: Mixed methods analysis of enterprise social networks. Comput. Networks. 75, 560–577 (2014)
18. Richter, A., Riemer, K.: The Contextual Nature of Enterprise Social Networking: A Multi Case Study Comparison. In: 21st European Conference on Information Systems, Utrecht, Paper 94 (2013)
19. Freeman, L.C.: Centrality in social networks conceptual clarification. Soc. Networks. 1, 215–239 (1978)
20. Gleave, E., Welser, H., Lento, T.M., Smith, M.A.: A Conceptual and Operational Definition of “Social Role” in Online Community. In: 42nd Hawaii International Conference on System Sciences. IEEE (2009)
21. Welser, H.T., Gleave, E., Fisher, D., Smith, M.A.: Visualizing the Signatures of Social Roles in Online Discussion Groups Finding Social Roles in Online Discussion. J. Soc. Struct. 8, 1–32 (2007)
22. Forestier, M., Stavrianou, A.: Roles in social networks: Methodologies and research issues. Web Intell. Agent Syst. 10, 117–133 (2012)
23. Trier, M., Richter, A.: The deep structure of organizational online networking - an actor-oriented case study. Inf. Syst. J. 25, 465–488 (2015)
24. Beck, R., Pahlke, I., Seebach, C.: Knowledge Exchange and Symbolic Action in Social Media-Enabled Electronic Networks of Practice: A Multilevel Perspective on Knowledge Seekers and Contributors. MIS Q. 38, 1245–1270 (2014)
25. Cetto, A., Klier, J., Klier, M., Richter, A., Wiesneth, K.: The Blessing of Giving: Kowledge Sharing and Knowledge Seeking in Enterprise Social Networks. In: 24th European Conference on Information Systems, Istanbul, paper 64 (2016)
26. Berger, K., Klier, J., Klier, M., Richter, A.: “who is Key...?” - Characterizing Value Adding Users in Enterprise Social Networks. In: 22nd European Conference on Information Systems, Tel Aviv, Paper 8 (2014)
27. Holtzblatt, L., Drury, J., Weiss, D.: Evaluating the Uses and Benefits of an Enterprise Social Media Platform. J. Soc. Media Organ. 1, 1–21 (2013)
28. AIS: Senior Scholars’ Basket of Journals, http://aisnet.org/?SeniorScholarBasket (Accessed: 25.08.2016)
29. VHB: VHB-JOURQUAL 3: Teilrating Wirtschaftsinformatik, http://vhbonline.org/en/service/jourqual/vhb-jourqual-3/teilrating-wi/ (Accessed: 25.08.2016)
30. Viol, J., Hess, J.: Information Systems Research on Enterprise Social Networks – A State-of-the-Art Analysis. In: Nissen, V., Stelzer, D., Straßburger, S., and Fischer, D. (eds.) Multikonferenz Wirtschaftsinformatik (MKWI) 2016, pp. 351–362. Universitätsverlag Ilmenau, Ilmenau (2016)

31. Kügler, M., Smolnik, S.: Uncovering the Phenomenon of Employees' Enterprise Social Software Use in the Post-Acceptance Stage – Proposing a Use Typology. In: 22nd European Conference on Information Systems, Tel Aviv, paper 1 (2014)
32. Mark, G., Guy, I., Kremer-Davidson, S., Jacovi, M.: Most liked, fewest friends. In: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14, pp. 393–404. ACM Press, New York (2014)
33. Steinhüser, M., Herzog, C., Peuker, V.: Nutzenpotenziale von Enterprise Social Software im Innovationsprozess. In: Nissen, V., Stelzer, D., Straßburger, S., and Fischer, D. (eds.) Multikonferenz Wirtschaftsinformatik (MKWI) 2016, pp. 339–350. Universitätsverlag Ilmenau, Ilmenau (2016)
34. Richter, A., Heidemann, J., Klier, M., Behrendt, S.: Success Measurement of Enterprise Social Networks. In: 11th International Conference on Wirtschaftsinformatik, Leipzig, paper 20 (2013)
35. Risius, M., Beck, R.: You Reap What You Sow? How Knowledge Exchange Effectiveness is Affected by Different Types of Communication in Enterprise Social Media. In: 47th Hawaii International Conference on System Sciences, pp. 540–549. IEEE (2014)
36. Riemer, K., Altenhofen, A., Richter, A.: What are you doing? - Enterprise Microblogging as Context Building. In: 19th European Conference on Information Systems, Helsinki, paper 252 (2011)
37. DiMicco, J.M., Millen, D.R., Geyer, W., Dugan, C., Brownholtz, B., Muller, M.: Motivations for Social Networking at Work. In: Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, pp. 711–720. ACM, New York (2008)
38. DiMicco, J.M., Geyer, W., Millen, D.R., Dugan, C., Brownholtz, B.: People Sensemaking and Relationship Building on an Enterprise Social Network Site. In: 42nd Hawaii International Conference on System Sciences, pp. 1–10. IEEE (2009)
39. van Osch, W., Steinfield, C.W.: Boundary Spanning through Enterprise Social Software: An External Stakeholder Perspective. In: 34th International Conference on Information Systems, Milan, paper 8 (2013)
40. Herzog, C., Richter, A.: Use Cases as a Means to Support the Appropriation of Enterprise Social Software. In: 2016 49th Hawaii International Conference on System Sciences (HICSS). pp. 4072–4081. IEEE (2016).
41. Raj, N., Dey, L., Gaonkar, B.: Expertise Prediction for Social Network Platforms to Encourage Knowledge Sharing. In: 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, pp. 380–383. IEEE (2011)
42. Wong, D., Bosua, R., Kurnia, S., Chang, S.: Exploring the Use of Enterprise 2.0 and Its Impact on Social Capital within a Large Organisation. In: ACIS 2015 Proceedings. pp. 1–10 (2015).
43. Seebach, C., Beck, R., Pahlke, I.: Situation Awareness Through Social Collaboration Platforms in Distributed Work Environments. In: 32nd International Conference on Information Systems, Shanghai, paper 2 (2011)
44. Richter, A., Riemer, K.: Corporate Social Networking Sites – Modes of Use and Appropriation through Co-Evolution. In: 20th Australasian Conference on Information Systems, Melbourne, paper 34 (2009)
45. Riemer, K., Diederich, S., Richter, A., Scifleet, P.: Short Message Discussions: On the Conversational Nature of Microblogging in a Large Consultancy Organisation. In: 15th Pacific Asia Conference on Information Systems, Brisbane, paper 158 (2011)
46. Burns, M.J., Kotval, X.P.: Questions About Questions: Investigating how Knowledge Workers Ask and Answer Questions. Bell Labs Tech. J. 17, 43–61 (2013)

Explizite Neutralität in Wählernetzwerken – Eine Analyse der Requests for Adminship (RfAs) in Wikipedia

Johannes Putzke[1], Hideaki Takeda[1]

[1] National Institute of Informatics, Principles of Informatics Research Division, Tokyo, Japan
{putzke,takeda}@nii.ac.jp

Abstract. Dieses Papier untersucht „Requests for Adminship (RfA)" in Wikipedia. RfAs sind Wahlen in Wikipedia, durch die Autoren zusätzliche technische Features erhalten können, die sie bei der Bearbeitung von Wikipedia unterstützen. Insbesondere beantworten wir die Forschungsfrage, was die Wahrscheinlichkeit erhöht, dass jemand eine neutrale Stimme gegenüber einem potentiellen Administrator abgibt. Zur Beantwortung der Forschungsfrage ziehen wir die Reziprozitätstheorie sowie die Balancetheorie heran. Die Ergebnisse weisen auf eine starke Tendenz zu neutraler Reziprozität (d. h. eine größere Wahrscheinlichkeit, dass ein Nutzer A eine neutrale Stimme für einen anderen Nutzer B abgibt, der wiederum auch eine neutrale Stimme für den ersten Nutzer A abgegeben hat) und neutraler Balance (d. h. eine größere Wahrscheinlichkeit, dass ein Nutzer A eine neutrale Stimme für einen anderen Nutzer B abgibt, der eine Gegenstimme von einem Nutzer C erhalten hat, für den der erste Nutzer A eine Gegenstimme abgegeben hat), hin.

Keywords: Analyse sozialer Netzwerke, Leadership, Online communities, Wiki, Wissensmanagement

1 Einleitung

Die Online-Enzyklopädie Wikipedia gehört zu den meist besuchten Webseiten der Welt[1] [1] und ist immer wieder Gegenstand wissenschaftlicher Forschungsarbeiten der Wirtschaftsinformatik bzw. ihrer anglo-amerikanischen Schwesterdisziplin *Information Systems* (IS) (z. B. [2-4]).

In diesem Kontext untersucht eine Forschungsströmung insbesondere Führungsverhalten in Wikipedia (z. B. [5]). Ein wichtiger Aspekt zum Verständnis von Führungsverhalten in Wikipedia sind die „*Requests for Adminship*" (RfAs). RfAs sind Wahlen in Wikipedia, durch die Autoren zusätzliche technische Features erhalten, die ihnen bei der Bearbeitung von Wikipedia behilflich sind und die ihnen eine stärkere Einflussnahme auf Inhalte der Wikipedia (etwa im Rahmen von Löschdiskussionen) ermöglichen. In einem öffentlichen, nicht-anonymisierten Wahlprozess können

[1] Vgl. http://www.alexa.com/siteinfo/wikipedia.org (Aufgerufen: 03.04.2016)

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Putzke, J.; Takeda, H. (2017): Explizite Neutralität in Wählernetzwerken - Eine Analyse der Requests for Adminship (RfAs) in Wikipedia, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 683-696

(registrierte) Wikipedia-Nutzer diese Ersuchen entweder unterstützen, gegen sie opponieren oder ihnen gegenüber eine neutrale Stimme[2] abgeben.

Obwohl eine Reihe von (hauptsächlich Konferenz-)Papieren RfAs im Detail untersucht (z. B. [5-13]), legt nach unserem Kenntnisstand keines dieser Arbeitspapiere einen Forschungsschwerpunkt auf die neutralen Stimmen bei diesen Wahlen. Vielmehr werden die neutralen Stimmen vor den weiteren Analysen ausgeschlossen (z. B. [8-10]), da sie als Störfaktoren angesehen werden.

Daher möchten wir in diesem Konferenzpapier tiefere Einsichten in die neutralen Stimmen der RfAs erhalten. Dies ist umso wichtiger, da ein substantieller Teil der Wikipedia-Nutzer (>5 Prozent) bei diesen Wahlen eine neutrale Stimme abgibt. Da diese Nutzer für ihre Stimmabgabe (Opportunitäts-)kosten investieren müssen, ist die Hauptforschungsfrage dieses Arbeitspapieres, weshalb werden neutrale Stimmen abgegeben? Da diese Forschungsfrage jedoch sehr allgemein gestellt ist und den Umfang eines Konferenzpapieres bei Weitem sprengen würde, möchten wir sie im Zugriff auf zwei Theorien – nämlich die Reziprozitätstheorie (z. B. [14, 15]) und die Balancetheorie (z. B. [16])(vgl. den nächsten Absatz) – näher spezifizieren. Insbesondere fragen wir uns, inwiefern diese beiden Theorien dazu geeignet sind, für den konkreten empirischen Fall die Frage zu beantworten:

F: Was erhöht die Wahrscheinlichkeit dafür, dass ein Nutzer eine neutrale Stimme für einen potentiellen Administrator abgibt?

Da nach unserem Kenntnisstand die beiden oben genannte Theorien noch nicht im Kontext von Wahlen getestet worden sind, insbesondere noch nicht im Kontext von Wahlen von Führungspersönlichkeiten in Online-Gemeinschaften, trägt unsere Forschung dazu bei, diese beiden Theorien auf diesen Kontext zu erweitern.

2 Theoretischer Hintergrund und Forschungshypothesen

Der Literaturüberblick und die Hypothesenentwicklung sind in zwei Teile strukturiert: Im ersten Teil wird ein Überblick über die verwandte Literatur bezüglich Wikis in der IS-Forschung gegeben, im zweiten Teil werden drei Forschungshypothesen entwickelt.

2.1 Wikis in der IS-Forschung

Neben den bereits in der Einleitung erwähnten Arbeiten zum RfA-Prozess steht dieses Papier in der Tradition von Papieren, die Wikis in der IS-Forschung analysieren. Wikis sind eine Zusammenstellung von Web-Seiten, die in einem kollaborativen Prozess

[2] Auch wenn neutrale Stimmen in Wikipedia zunächst keinen Einfluss auf Wahlen zu nehmen scheinen, können sie durchaus wahlbeeinflussend wirken. Das Ergebnis einer Wahl entscheidet sich aus dem Verhältnis der unterstützenden Stimmen zu den opponierenden Stimmen. Obwohl die neutralen Stimmen bei diesen Berechnungen keinen direkten Eingang finden, werden die neutralen Stimmen in Wikipedia von den jeweiligen Wählern mit Kommentaren / Begründungen für ihre neutrale Stimmabgabe versehen. In der Konsensfindung, ob ein Kandidat in den Admin-Status erhoben werden soll, werden auch diese qualitativen Kommentare berücksichtigt.

erstellt und iterativ verändert werden [17]. In diesem Kontext sind insbesondere zwei Forschungsströmungen für dieses Papier relevant. Die erste Forschungsströmung untersucht Wikis aus einer Wissensmanagementperspektive (z. B. [17, 18]), die zweite Forschungsströmung untersucht Führungsverhalten in Wikis (z. B. [1, 19]).

Die erste Forschungsströmung ist tief in der Wissensmanagement-Literatur verankert (z. B. [20-22]). Ein guter Literaturüberblick über diese Forschung findet sich in Beck, Rai, Fischbach und Keil [18].

In dieser Forschungstradition wird eine wesentliche Unterscheidung zwischen zwei verschiedenen Arten von Beiträgen auf Wikipedia gegeben: Wissenserschaffung („*knowledge adding*") und Wissensformung („*knowledge shaping*"). Wissensformung bezeichnet dabei die kontinuierliche Revision von Beiträgen zu einem Wiki [17]. Auch Beck, Rai, Fischbach und Keil [18] folgen dieser Unterscheidung und sprechen von „*knowledge creation*" und „*knowledge integration*". Wissensintegration definieren sie dabei als die Reorganisation, Verknüpfung und Synthese von Wissen, das von anderen beigetragen wurde. Während eine Reihe von Papieren die Wissenserschaffung im Detail untersucht, besteht hinsichtlich der Untersuchung von Wissensformung noch eine Forschungslücke. In diesem Zusammenhang rufen IS-Forscher insbesondere zu Forschung auf, die das Teilen von Wissen in Wikis aus der Perspektive der Analyse sozialer Netzwerke heraus analysiert [18]. Mit diesem Beitrag antworten wir auf diese Aufforderung.

Die zweite Forschungsströmung untersucht Führungsverhalten in Wikis (z. B. [1, 19]). Da die Übertragbarkeit traditioneller Theorien des Führungsverhaltens auf Online-Führungsverhalten aufgrund mediierter Kommunikation, mangelnden persönlichen Interaktionen, der Vorherrschaft text-basierter, asynchroner Nachrichten sowie einer „bottom-up" Führungsstruktur fraglich ist, rufen Forscher zu Theorien des Online-Führungsverhalten auf [23]. Mit diesem Papier antworten wir auch auf diese Aufforderung. Dabei erweitern wir die Arbeiten von Forschern, die den RfA-Prozess von Wikipedia im Detail untersuchen (z. B. [5-13]), indem wir die vorhergenannten theoretischen Perspektiven durch zwei interdisziplinäre Theorien anreichern, die in der Hypothesenentwicklung näher beleuchtet werden: Die „Reziprozitätstheorie" sowie die „Balancetheorie".

2.2 Forschungshypothesenentwicklung

Die erste und zweite Hypothese greifen auf die Reziprozitätstheorie zurück (z. B. [14, 15, 24])). In diesem Papier definieren wir Reziprozität als das Zurückgeben von etwas Äquivalentem. Viele Redewendungen wie etwa „wie Du mir, so ich Dir", „Gleiches mit Gleichem vergelten" oder „quid pro quo" spiegeln die Norm der Reziprozität in menschlichen Gesellschaften wider. Dabei umfasst die Norm der Reziprozität eine positive Valenz („Eine Hand wäscht die andere") ebenso wie eine negative („Auge um Auge, Zahn um Zahn").

Während die Reziprozitätstheorie explizit indirekte Formen der Reziprozität mit einschließt (für einen Literaturüberblick vgl. [25]), fokussieren wir in diesem Papier auf direkte Reziprozität. (Indirekte Reziprozität findet sich etwa im Prinzip "I scratch your back and someone else will scratch mine" [25]; vgl. in diesem Kontext auch Faraj und Johnson [26]). Insbesondere konzeptionalisieren wir Reziprozität im Sinne der Netzwerktheorie, d. h. in diesem Papier beschreibt Reziprozität die erhöhte Wahrscheinlichkeit, dass ein Akteur A eine Kante zu denjenigen Akteuren aufbaut, die wiederum eine Kante zu Akteur A aufgebaut haben [27].

In der Netzwerktheorie wurde Reziprozität insbesondere für Netzwerke mit einer positiven Valenz (wie etwa Freundschaftsnetzwerke) untersucht (für eine Ausnahme vgl. Rambaran, Dijkstra, Munniksma und Cillessen [28]). Der Grund hierfür mag der Tatsache zuzuschreiben sein, dass die Basisrate, sich von einer asymmetrischen Beziehung zu einer reziprok-symmetrischen Beziehung zu entwickeln, in Netzwerken mit einer negativen Valenz (wie etwa Feindschaftsnetzwerken) niedriger ist als in Netzwerken mit einer positiven Valenz [29].[3]

Für Reziprozität in Netzwerken mit einer positiven Valenz (positive Reziprozität) gibt es eine Reihe theoretischer Erklärungen (vgl. [30]). Die bekanntesten stammen aus der Theorie sozialer Austausche (z. B. [31, 32]). Da die Hypothese der positiven Reziprozität jedoch bereits ausführlich hergeleitet und getestet worden ist (z. B. [33, 34]), soll sie in diesem Papier nicht näher behandelt werden.

Weniger Aufmerksamkeit erfahren hat hingegen die Hypothese der negativen Reziprozität. Nichtsdestotrotz ist das Konzept der negativen Reziprozität (Rache) tief in unserer Gesellschaft verankert und findet sich auch in frühen philosophischen Schriften wieder (vgl. [35]). Daher ist es angemessen zu erwarten, dass Akteure unwohlwollendes Verhalten anderer ihnen gegenüber erwidern [35]. Im Fall der RfAs ist eine Gegenstimme für einen Nutzer, der als Administrator kandidiert, ein Beispiel für solch ein unwohlwollendes Verhalten. Daher stellen wir die Hypothese auf

> *H1 (negative Reziprozität). Die Wahrscheinlichkeit dass Nutzer A eine Gegenstimme für Nutzer B abgibt, der als Administrator kandidiert, ist größer, wenn Nutzer B zuvor auch eine Gegenstimme für Nutzer A abgegeben hat.*

Während Für- und Gegenstimmen in RfAs eine klare Valenz/Wertigkeit haben, ist die Valenz neutraler Stimmen weniger eindeutig. Während neutrale Stimmen in Wahlen, in denen die Nutzer größtenteils mit Gegenstimmen votiert haben, eine positive Valenz haben, haben Sie in Wahlen, in denen die Nutzer größtenteils mit Für-Stimmen votiert haben, eine negative Valenz. Nichtsdestotrotz nehmen wir nach der Reziprozitätstheorie an, dass Nutzer eine neutrale Stimme gleich ihrer Valenz erwidern und stellen die Hypothese auf[4]

---

[3] Auf der anderen Seite sei darauf hingewiesen, dass es zwischen Forschern einen emergenten Konsens gibt, dass die Neigung zur Bestrafung unwohlwollendes Verhaltens größer ist, als die Neigung wohlwollendes Verhalten zu belohnen [15].

[4] An dieser Stelle mag sich der Leser fragen, wo eine ursprünglich neutrale Stimme, die erwidert wird, ihren Ursprung hat. Hierfür gibt es eine Reihe von Erklärungen, die den Rahmen dieses Papiers sprengen würden. Zum Beispiel gewinnt ein User (nach der „Signaling-Theorie”) Reputation als aktiver Bewerter. Außerdem könnte der User dadurch Satisfaktion erfahren, dass er zu einem öffentlichen Gut beiträgt [36].

H2 (neutrale Reziprozität). Die Wahrscheinlichkeit, dass Nutzer A eine neutrale Stimme für Nutzer B abgibt, der als Administrator kandidiert, ist größer, wenn Nutzer B zuvor auch eine neutrale Stimme für Nutzer A abgegeben hat.

Die Hypothesenentwicklung der dritten Hypothese basiert auf der Balancetheorie. Die Ursprünge der Balancetheorie liegen mehr als ein halbes Jahrhundert zurück (für einen Literaturüberblick über die Balancetheorie in der IS-Forschung vgl. [16]). Die Balancetheorie wurde zunächst von Heider [37] in die psychologische Literatur eingeführt und anschließend von Cartwright und Harary [38] verallgemeinert, die auch eine formale Definition von „Balance“ im Rahmen der Graphentheorie liefern.

Die fundamentalen Grundsätze der (strukturellen) Balancetheorie sind: a) Die Freunde meines Freundes sind meine Freunde, b) die Feinde meines Freundes sind meine Feinde, c) die Freunde meines Feindes sind meine Feinde und d) die Feinde meines Feindes sind meine Freunde. Während diese fundamentalen Grundsätze der Balancetheorie auf zwei Zustände dyadischer Beziehungen fokussieren (d.h. Freundschaft und Feindschaft), erweitern wir die strukturelle Balancetheorie in diesem Papier durch einen dritten Zustand dyadischer Beziehungen. Insbesondere lassen wir „explizite Neutralität“ gegenüber einem anderen Akteurs im Wahlnetzwerk zu. D. h. wir machen keine so starke Proposition, dass Feinde eines Feindes Freunde werden, sondern wir nehmen an, dass jemand zumindest eine neutrale Stimme gegenüber dem „Feind“ eines „Feindes“ abgeben sollte[5] und stellen die Hypothese auf:

H3 (neutrale Balance): Die Wahrscheinlichkeit, dass Nutzer A eine neutrale Stimme für Nutzer B abgibt, der eine Gegenstimme von Nutzer C erhalten hat, für den Nutzer A wiederum mit einer Gegenstimme votiert hat, ist größer als die Wahrscheinlichkeit, dass Nutzer A eine zufällige neutrale Stimme abgibt.

3 Methode

3.1 Datensammlung und Stichprobe

Um die Hypothesen zu testen, benutzen dwir einen Datensatz von Leskovec, Huttenlocher und Kleinberg [39]. Dieser Datensatz enthält alle Wahlergebnisse der „Requests for Adminship” (RfA) in Wikipedia bis Januar 2008. Der gleiche Datensatz wurde von einer Reihe von Forschern benutzt. Auch wenn dieser Datensatz vergleichsweise alt ist, bringt dessen Verwendung den Vorteil, zu vorhergehender Forschung vergleichbar zu sein [12].

Insgesamt enthält dieser Datensatz die Wahlergebnisse von 2,794 Wahlen mit 114,040 Stimmen (von diesen 83,962 Für-Stimmen, 23,118 Gegenstimmen, und 6,960 neutrale Stimmen). 11 der Stimmen hatten einen Zeitstempel mit einer Jahreszahl > 2050 und wurden deshalb von den weiteren Analysen ausgeschlossen.

---

[5] Ein „Feind” im Wahlnetzwerk eines Nutzers A ist ein Wähler, der während Nutzer As RfA eine Gegenstimme abgegeben hat.

Außerdem schlossen wir vor den weiteren Analysen alle Stimmen mit einem Zeitstempel größer als 31-12-2007 und kleiner als 1-1-2005 aus. Diese Entscheidung wurde getroffen, da wir den Datensatz für die weiteren Analysen in äquidistante Zeitperioden aufteilen wollten (siehe das nächste Unterkapitel). Da die Wahldaten in 2004 und 2008 nicht sehr umfangreich sind (z. B. gibt es im Januar 2008 nur 6 Tage, für die Wahlergebnisse vorliegen), ist es zweckreich die Daten auf ihrer jährlichen Basis zu untersuchen, für die die komplette Stimmhistorie vorliegt (d. h. 2005-2007). Insgesamt stellten sich – nach Ausschluss der oben genannten Stimmen – 1972 verschiedene Akteure zur Wahl und ließen sich von 3219 verschiedenen Akteuren wählen. Da einige der Akteure sowohl als Gewählte als auch als Wählende in Erscheinung treten, basieren die folgenden Analysen der RfAs auf einem Datensatz mit insgesamt 4,327 verschiedenen Akteuren.[6]

3.2 Modell

Zum Hypothesentest benutzten wir ein stochastisches, akteurbasiertes Modell, das von Snijders in der soziologischen und statistischen Literatur vorgeschlagen wurde [40-42]. Nach unserem Kenntnisstand ist die erste Anwendung dieser Methode in der IS-Forschung ein Papier von Putzke, Fischbach, Schoder und Gloor [43]. Eine volle Beschreibung des Modellierungsansatzes würde den Rahmen dieses Papiers bei Weitem sprengen. Der interessierte Leser sei daher auf die eben genannte Literatur verwiesen. In den folgenden Absätzen stellen wir lediglich die wichtigsten Punkte dieses Modellierungsansatzes vor, so dass dem Leser ein Verständnis des Ergebniskapitels und die Reproduzierbarkeit der Forschungsergebnisse ermöglicht werden.

Im Folgenden bezeichnet $X(t)=X_{ij}(t)$ eine $n\times n$ $(n=4,327)$ Adjazenzmatrix des Netzwerkes neutraler Stimmen, wobei $X_{ij}=1(0)$ eine Kante (keine Kante) von Akteur i zu Akteur j $(i,j=1,\ldots n)$ in Periode t repräsentiert, d. h., Nutzer i gab bis Periode t eine neutrale Stimme für Nutzer j $(i\rightarrow j)$ ab. Analog bezeichnet $W(t)=W_{ij}(t)$ eine $n\times n$ Adjazenzmatrix des Netzwerkes der Gegenstimmen, wobei $W_{ij}=1(0)$ eine Kante (keine Kante) von Akteur i zu Akteur j $(i,j=1,\ldots n)$ in Periode t repräsentiert, d. h., Nutzer i gab eine Gegenstimme für Nutzer j $(i\rightarrow j)$ ab.

Das Modell nimmt an, dass diese Adjazenzmatritzen eine Markovkette in kontinuierlicher Zeit mit stationären Übergangsmatritzen bilden. Um die Übergangsmatritzen dieser Markovkette herzuleiten, zerlegt Snijders [40-42] die Veränderung zwischen zwei Zuständen der Markovkette in zwei grundlegende Modelle: (1) Ein Wartezeitmodell für Kantenveränderungen und 2) ein Modell für die Wahl der Kantenveränderung.

Das Wartezeitmodell für die Kantenveränderungen nimmt an, dass in zufälligen, deterministischen Momenten der Zeit (so genannten „Minischritten", *engl.*

[6] An dieser Stelle sei erwähnt, dass in den RfAs eine Reihe von Nutzern mehrfach kandidiert. Aus Sicht des Kandidaten gibt es hierfür hauptsächlich zwei Gründe: Erstens kann ein Nutzer durch eine Wiederwahl eine höhere Legitimität seines Admin-Status erhalten. Zweitens versuchen Nutzer durch eine zweite Wahl Admin-Status zu erhalten, wenn sie in der ersten Wahl nicht erfolgreich waren.

„ministeps“) die Akteure die Möglichkeit erhalten, eine Kantenvariable zu verändern. Dabei wird angenommen, dass die Wartezeit zwischen zwei dieser Momente einer Exponentialverteilung folgt. In Tabelle 1 im Ergebniskapitel stehen diese Ratenfunktionen in den Zeilen „Änderungsrate des Netzwerkes“. Dabei nehmen wir an dieser Stelle keine konstante Änderungsrate des Netzwerkes für jede Periode an, sondern gehen davon aus, dass die Änderungsrate des Netzwerkes von Periode zu Periode variiert. (Eine Periode in diesem Papier entspricht dabei einem Jahr). In den Zeilen „Änderungsraten des Netzwerks“ stehen also die Parameter λ der Exponentialverteilungen, die die Wartezeit auf einen Minischritt in den jeweiligen Perioden beschreiben.

Das Modell für die Wahl der Kantenveränderungen ist durch sogenannte „Evaluationsfunktionen” (*engl. „evaluation functions”*) repräsentiert. Akteur is Evaluationsfunktion für das Netzwerk neutraler Stimmen ist definiert als $f_i^X(x,w) = \sum_k \beta_k^X s_{ki}^X(x,w)$ und Akteur is Evaluationsfunktion für das Netzwerk der Gegenstimmen (in ihrer generellen Form) ist definiert als als $f_i^W(x,w) = \sum_k \beta_k^W s_{ki}^W(x,w)$, wobei die $s_{ki}(x,w)$ so genannte Effekte (*engl. „effects*”) sind. Diese Effekte spiegeln Akteur is Einbettung in diese Netzwerke wider. Die *βs* der Evaluationsfunktionen können wie (logistische Regressionskoeffizienten) interpretiert werden. Daher wurden für die Hypothesentests die oben genannten Effekte wie folgt spezifiziert:

Negative Reziprozität wurde als $s_{1i}^W = \sum_j w_{ij} w_{ji}$ gemessen. Das heißt, Nutzer is Evaluationsfunktion steigt nur dann um den Wert 1, wenn Nutzer *i* eine Gegenstimme für Nutzer *j* abgibt ($w_{ij} = 1$) und Nutzer *j* eine Gegenstimme für Nutzer *i* abgibt ($w_{ji} = 1$). Wenn eine der beiden Stimmen fehlt (d. h. $w_{ij} = 0$ oder $w_{ji} = 0$), ist das Produkt der beiden Terme 0. Folglich zeigt ein positiver Parameter β_1^W eine größere Wahrscheinlichkeit dafür an, dass Nutzer *i* eine Gegenstimme für Nutzer *j* abgibt, wenn Nutzer *j* auch eine Gegenstimme für Nutzer *i* abgibt. Der Schätzer für den Parameter β_1^W findet sich in der Zeile „Negative Reziprozität” in Tabelle 1.

Neutrale Reziprozität wurde als $s_{1i}^X = \sum_j x_{ij} x_{ji}$ gemessen. Die Erklärung für diesen Effekt ist analog zu derjenigen der negativen Reziprozität. Wiederum zeigt ein positiver Parameter β_1^X eine größere Wahrscheinlichkeit an, dass Nutzer *i* eine neutrale Stimme für Nutzer *j* abgibt, wenn Nutzer *j* auch eine neutrale Stimme für Nutzer *i* abgibt. Der Schätzer für den Parameter β_1^X findet sich in der Reihe „neutrale Reziprozität” in Tabelle 1.

Schließlich wurde, *neutrale Balance* als $s_{2i}^X = \sum_{j \neq h} x_{ij} w_{ih} w_{hj}$ gemessen (für eine Erklärung vgl. oben). Der Schätzer für Parameter β_2^X findet sich in der Zeile „neutrale Balance” in Tabelle 1.

4 Ergebnisse

Das vorgeschlagene Modell wurde in R 3.2.2 unter Verwendung des Pakets RSiena (*„Simulation Investigation for Empirical Network Analysis“*), Version 1.1-290 [44] geschätzt. Für die Schätzung wurde die Momentenmethode mit Robbins-Monro [45] stochastischem Approximationsalgorithmus verwendet. T-Tests deuten auf eine gute

Konvergenz des Modells hin (angewendetes Kriterium: $t < .1$), ebenso wie das globale maximale Konvergenzverhältnis $t_{convmax}= 0.1009$ (angewendetes Kriterium: $t_{convmax} < .25$) (vgl. [46]).

Tabelle 1 veranschaulicht die Ergebnisse der Modellschätzung. (Tabelle 2 im Anhang enthält die Korrelationskoeffizienten zwischen den Parameterschätzern. Es gab keine Hinweise auf Multikollinearität.) Alle Effekte des Modells sind statistisch höchst signifikant. Daher werden Hypothesen H1-H3 unterstützt. Es gibt eine starke Tendenz für negative Reziprozität im Netzwerk der Gegenstimmen, was bedeutet, dass die Wahrscheinlichkeit größer ist, das Nutzer A eine Gegenstimme für einen Nutzer B abgibt, wenn dieser Nutzer B auch eine Gegenstimme für den ersten Nutzer A abgegeben hat (H1). Des Weiteren ist die Wahrscheinlichkeit größer, dass ein Nutzer A eine neutrale Stimme für einen Nutzer B abgibt, wenn Nutzer B zuvor auch mit einer neutralen Stimme für Nutzer A votiert hat (H2). Schließlich ist die Wahrscheinlichkeit größer, dass ein Nutzer A eine neutrale Stimme für einen Nutzer B abgibt, der eine Gegenstimme von einem Nutzer C erhalten hat, für den wiederum der erste Nutzer A eine negative Stimme abgegeben hat, als eine zufällige neutrale Stimme abzugeben (H3).

Tabelle 1. Modellergebnisse

	par.	*s.e.*	*t-value*	*p-value*
Netzwerk der Gegenstimmen				
Änderungsrate des Netzwerks (t=1)	2.213	0.0223	99.238	0.0000***
Änderungsrate des Netzwerks (t=2)	2.082	0.0217	95.949	0.0000***
Negative Reziprozität (H1)	1.605	0.1296	12.387	0.0000***
Netzwerk der neutralen Stimmen				
Änderungsrate des Netzwerks (t=1)	0.681	0.0126	54.048	0.0000***
Änderungsrate des Netzwerks (t=2)	0.614	0.0117	52.504	0.0000***
Neutrale Reziprozität (H2)	2.181	0.3267	6.676	0.0000***
Neutrale Balance (H3)	0.601	0.1285	4.677	0.0000***

Legend: *par.* = Parameterschätzer (*λ/β*) , *s.e.* = Standardfehler
*** p<0.001, ** p<0.01, * p<0.05

An dieser Stelle mag sich der Leser die Frage nach der Effektgröße der Parameterschätzer stellen. Wie bereits oben erwähnt können die Parameterschätzer wie logistische Regressionskoeffizienten interpretiert werden. Daher ist es auch möglich, das Chancenverhältnis (*engl. odds ratio*) der unabhängigen Variablen zu berechnen. In diesem Fall kann das Chancenverhältnis als der Faktor interpretiert werden, durch den sich das Verhältnis zwischen der Wahrscheinlichkeit, während eines Minischrittes eine Stimme abzugeben, und der Wahrscheinlichkeit, keine Stimme abzugeben, ändert, wenn sich die im Fokus stehende unabhängige Variable um eine Einheit erhöht und alle anderen unabhängigen Variablen konstant bleiben. Das Chancenverhältnis lässt sich berechnen als $e^{\beta_k^{X,W}}$.

Zum Beispiel ändert sich bezüglich H1 das Verhältnis zwischen der Wahrscheinlichkeit während eines Ministeps eine Gegenstimme abzugeben und der Wahrscheinlichkeit keine Gegenstimme abzugeben um den Faktor $e^{1.605} = 4.978$, wenn der Wähler mit seiner Stimme während dieses Minischritts eine Gegenstimme erwidert. Vergleichbar steigt das Verhältnis zwischen der Wahrscheinlichkeit, während eines Minischrittes eine neutrale Stimme abzugeben, und der Wahrscheinlichkeit, keine neutrale Stimme abzugeben, um den Faktor $e^{2.181} = 8.855$, wenn der Wähler mit seiner Stimme eine neutrale Stimme erwidert (H2). Schließlich beträgt das Chancenverhältnis in Bezug auf neutrale Balance $e^{0.601} = 1.824$ (H3). Wiederum werden die Forschungshypothesen stark durch die Chancenverhältnisse unterstützt.

5 Diskussion

Das Hauptanliegen dieses Papiers war es herauszufinden, was die Wahrscheinlichkeit erhöht, dass jemand eine neutrale Stimme für einen potentiellen Administrator in Wikipedia abgibt. Die Ergebnisse liefern starke Evidenz für neutrale Reziprozität (d. h. eine höhere Wahrscheinlichkeit, dass ein Nutzer A eine neutrale Stimme für einen Nutzer B abgibt, der auch eine neutrale Stimme für den ersten Nutzer A abgegeben hat) sowie für neutrale Balance (d. h. eine größere Wahrscheinlichkeit, dass ein Nutzer A eine neutrale Stimme für einen anderen Nutzer B abgibt, der eine Gegenstimme von einem Nutzer C erhält, für den der erste Nutzer A mit einer Gegenstimme votiert hat).

5.1 Theoretischer Beitrag und Implikationen

Dieses Papier liefert mindestens zwei theoretische Beiträge. Erstens trägt es zu unserem Verständnis der RfAs bei. Während die meisten Studien über RfAs deskriptive Studien sind, testet dieses Papier formal drei Hypothesen über RfAs. Hierbei wendet das Papier die Reziprozitätstheorie (z. B. [14, 15]) und die Balancetheorie (z. B. [16]) in einem neuen Kontext an (konkret im Kontext der Wahlen von Führern von Online Communities). Die Ergebnisse zeigen, dass die Propositionen beider Theorien auch in diesem neuen Kontext Bestand haben. Dies ist keinesfalls selbstverständlich, da sich Interaktionen in Online Communities substantiell von offline Interaktionen unterscheiden. Beispiele für solche Differenzen sind etwa mediierte Kommunikation und die Vorherrschaft text-basierter, asynchroner Kommunikation. Zum Beispiel beeinflusst die asynchrone, textbasierte Kommunikation die vermuteten Kausalverhältnisse in zwei Richtungen. Zum einen ist zu erwarten, dass durch die schriftliche Fixierung einer Stimmabgabe diese den entsprechenden Akteuren länger in Erinnerung bzw. abrufbar bleibt. Daher sollten die entsprechenden Kausalverhältnisse in Online Communities stärker ausgeprägt sein als bei mündlicher, synchroner Kommunikation. Auf der anderen Seite bedeutet asynchrone Kommunikation jedoch auch eine Zeitverzögerung, die sich negativ auf die vermuteten Kausalverhältnisse auswirken sollte. Zukünftige Forschung sollte daher auch entsprechende Hypothesen testen, die genau auf diese Unterschiede abzielen.

Zweitens wurde die Balancetheorie um einen neuen Zustand dyadischer Beziehungen erweitert. Konkret wurden die beiden Zustände einer positiven Beziehung und einer negativen Beziehung durch den Zustand einer „*expliziten Neutralität*“ erweitert.

Dennoch stellt sich an dieser Stelle die Frage nach der Übertragbarkeit der Studienergebnisse auf Wahlen im Allgemeinen. RfAs haben mehrere Charakteristika, die eine einfache Übertragbarkeit der Studienergebnisse auf Wahlen im Allgemeinen in Frage stellen. Erstens sind RfAs nicht geheim. Zweitens sind neutrale Stimmen in Wikipedia mit einem Kommentar bzw. einer Begründung für die Stimmabgabe versehen. Drittens wird jede Stimme und deren Begründung mit einem Zeitstempel dokumentiert. Während längst nicht jede Wahl geheim ist, sind die sukzessive Stimmabgabe und deren Dokumentation durchaus als besondere Merkmale der RfAs zu begreifen. Dennoch denken wir, dass wir mit diesem Papier dadurch einen theoretischen Beitrag leisten, dass wir die Anwendbarkeit der Reziprozitätstheorie und der Balancetheorie in einem neuen, wenn auch speziellen Kontext zeigen.

In Paragraph 2.1 haben wir das große Interesse der Forschungsgemeinschaft an diesem speziellen Kontext bereits aufgezeigt. Insbesondere hat die Forschungsgemeinschaft auch hier ein großes Interesse daran gezeigt, die RfAs aus einer Perspektive der Analyse sozialer Netzwerke heraus zu analysieren (vgl. [5-12]). Dennoch geht dieses Forschungspapier mit der formalen Herleitung seiner Hypothesen und deren Test weit über diese Arbeiten hinaus. Als neu dabei ist hier vor allem die Fokussierung auf neutrale Stimmen in den Wahlen zu sehen, die in vorherigen Analysen sogar oftmals ausgeschlossen worden sind.

5.2 Limitationen und zukünftige Forschung

Wie jede empirische Studie unterliegt auch diese Studie Limitationen, die ihre Relevanz und Rigorität (*engl. „rigor“*) einschränken. Wir denken nicht, dass diese Limitationen die Ergebnisse obsolet machen, so lange sich der Leser derer bewusst ist, während er seine Schlussfolgerungen zieht. Vielmehr bieten sie Raum für zukünftige Forschung und zur Weiterentwicklung dieses Papiers in einen Fachzeitschriftenartikel.

Erstens wird keine Maßzahl für die Anpassungsgüte des Modells berichtet. Jedoch existiert für die Klasse der stochastischen, akteurs-orientierten Modelle bisher kein globales Anpassungsgütemaß wie etwa R^2 in der linearen Regression (z. B. [47]). Eine Alternative könnte es sein, die Modellgüte durch Score-Tests [48] oder einen Monte-Carlo-Mahalanobis-Distanz-Test auf Basis einer Hilfsstatistik (wie etwa der „*indegree distribution*“) zu bestimmen.[7] Unglücklicherweise ist das untersuchte Netzwerk jedoch aufgrund der Rechenzeiten zu groß, um diese Tests auf handelsüblichen PCs durchzuführen. Diese Tests wären jedoch ein lohnendes Unterfangen für zukünftige Forschung.

[7] Diese Art von Test wurde von Lospinoso vorgeschlagen. Nach unserem Kenntnisstand wurde bisher kein Artikel mit einer fundamentalen Beschreibung dieser Methode publiziert. Der interessierte Leser findet eine Präsentation über diese Art von Test unter http://www.stats.ox.ac.uk/~snijders/siena/SienaGOF_s.pdf (Aufgerufen: 07.04.2016). Eine Anwendung dieses Tests in der IS Forschung findet sich in [49].

Zweitens wurden keine Robustheits-Checks für das Modell ausgeführt. Zukünftige Forschung sollte das Modell zum Beispiel auch mit einigen Kontrollvariablen schätzen.

Drittens wurde der Datensatz für die Analyse in drei äquidistante Zeitperioden aufgespalten, wodurch Zeit-Informationen verloren gingen. Zukünftige Forschung könnte daher darüber nachdenken, die Daten als Event-Daten zu modellieren.

Viertens wurden in unseren Analysen die Fürstimmen vernachlässigt. Auch wenn diese in der bisherigen Forschung bereits intensiv untersucht worden sind, könnte die zukünftige Forschung das Zusammenspiel von neutralen, positiven und negativen Stimmen im Detail untersuchen. Als ersten Schritt in diese Richtung untersuchten wir, ob positive oder neutrale Stimmen nach negativen Stimmen weniger wahrscheinlich sind und ob positive oder negative Stimmen nach neutralen Stimmen weniger wahrscheinlich sind. In Tabelle 3 finden sich in den ersten drei Zeilen die unbedingten Wahrscheinlichkeiten für eine negative, neutrale und positive Stimme (ohne die Anfangsstimmen der Wahl) als Vergleichswert. In den übrigen Zeilen der Tabelle stehen die bedingten Wahrscheinlichkeiten für eine negative (neutrale, positive) Stimme unter der Bedingung, dass die Vorstimme negativ (neutral, positiv) war. Auch diese Analysen stützen unsere Ergebnisse und zeigen die Bedeutung von neutralen Stimmen für Wahlprozesse in Wikipedia auf.

Fünftens wählten wir mit nicht-anonymen Online-Wahlen von Führern in einer virtuellen Gemeinschaft einen sehr spezifischen Rahmen. Daher besteht nach wie vor die Frage, ob die Ergebnisse dieser Studie (neben den oben genannten besonderen Merkmalen der RfAs) auch auf Wahlen im Allgemeinen übertragbar sind. Zukünftige Forschung sollte diesen Punkt im Detail untersuchen.

Tabelle 3. Wahrscheinlichkeiten für eine negative, neutrale und positive Stimme

	# Stimmen	*Wahrscheinlichkeit in %*
Unbedingte Wahrscheinlichkeiten		
Negative Stimme	22332	20.0744
Neutrale Stimme	6791	6.1045
Positive Stimme	82123	73.8211
Bedingte Wahrscheinlichkeiten		
Negative Stimme **|** *Vorstimme negativ*	11237	50.9153
Neutrale Stimme **|** *Vorstimme negativ*	2632	11.9257
Positive Stimme **|** *Vorstimme negativ*	8201	37.1590
Negative Stimme **|** *Vorstimme neutral*	2493	37.0982
Neutrale Stimme **|** *Vorstimme neutral*	993	14.7768
Positive Stimme **|** *Vorstimme neutral*	3234	48.1250
Negative Stimme **|** *Vorstimme positiv*	8602	10.4322
Neutrale Stimme **|** *Vorstimme positiv*	3166	3.8396
Positive Stimme **|** *Vorstimme positiv*	70688	85.7281

Anhang

Tabelle 2. Korrelationsmatrix der Schätzer

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Netzwerk der Gegenstimmen							
(1) *Änderungsr. d. Netzw.(t=1)*	1						
(2) *Änderungsr. d. Netzw.(t=2)*	-.03	1					
(3) *Negative Reziprozität*	-.025	-.029	1				
Netzwerk der neutr. Stimmen							
(4) *Änderungsr. d. Netzw.(t=1)*	-.043	-.017	.073	1			
(5) *Änderungsr. d. Netzw.(t=2)*	-.049	-.007	-.008	-.04	1		
(6) *Neutrale Reziprozität*	.027	.044	.013	-.027	-.002	1	
(7) *Neutrale Balance*	.056	-.028	.024	-.021	-.077	.03	1

Danksagungen

Die Arbeit wurde mit Unterstützung eines Stipendiums im Rahmen des FITweltweit-Programms des DAAD ermöglicht.

References

1. Zhu, H.Y., Kraut, R.E., Kittur, A.: Effectiveness of Shared Leadership in Wikipedia. Hum Factors 55, 1021-1043 (2013)
2. Ransbotham, S., Kane, G.C.: Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in Wikipedia. Mis Quart 35, 613-627 (2011)
3. Xu, S.X., Zhang, X.Q.: Impact of Wikipedia on Market Information Environment: Evidence on Management Disclosure and Investor Reaction. Mis Quart 37, 1043-1068 (2013)
4. Kane, G.C., Johnson, J., Majchrzak, A.: Emergent Life Cycle: The Tension Between Knowledge Change and Knowledge Retention in Open Online Coproduction Communities. Manage Sci 60, 3026-3048 (2014)
5. Collier, B., Burke, M., Kittur, N., Kraut, R.: Promoting Good Management: Governance, Promotion, and Leadership in Open Collaboration Communities. In: International Conference on Information Systems (ICIS), pp. 220. (2010)
6. Cabunducan, G., Castillo, R., Lee, J.B.: Voting Behavior Analysis in the Election of Wikipedia Admins. In: 2011 International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 545-547. (2011)
7. Jankowski-Lorek, M., Ostrowski, L., Turek, P., Wierzbicki, A.: Modeling Wikipedia Admin Elections Using Multidimensional Behavioral Social Networks. Social Network Analysis and Mining 3, 787-801 (2013)
8. Lee, J.B., Cabunducan, G., Cabarle, F.G.C., Castillo, R., Malinao, J.A.: Uncovering the Social Dynamics of Online Elections. J Univers Comput Sci 18, 487-506 (2012)

9. Leskovec, J., Huttenlocher, D.P., Kleinberg, J.M.: Governance in Social Media: A Case Study of the Wikipedia Promotion Process. In: Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, pp. 98-105. (2010)
10. Porco, A., Kaltenbrunner, A., Gómez, V.: Low-rank Approximations for Predicting Voting Behaviour. Workshop on Networks in the Social and Information Sciences, NIPS 2015, Montreal, Quebec, Canada (2015)
11. West, R., Paskov, H.S., Leskovec, J., Potts, C.: Exploiting Social Network Structure for Person-to-Person Sentiment Analysis, http://arxiv.org/abs/1409.2450 (Aufgerufen: 03.04.2016)
12. Picot-Clemente, R., Bothorel, C., Jullien, N.: Contribution, Social networking, and the Request for Adminship process in Wikipedia. In: OpenSym 2015 : 11th International Symposium on Open Collaboration. ACM (2015)
13. Burke, M., Kraut, R.: Mopping Up: Modeling Wikipedia Promotion Decisions. Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, pp. 27-36. ACM, San Diego, CA, USA (2008)
14. Falk, A., Fischbacher, U.: A theory of reciprocity. Games and Economic Behavior 54, 293-315 (2006)
15. Fehr, E., Gächter, S.: Fairness and retaliation: The economics of reciprocity. J Econ Perspect 14, 159-181 (2000)
16. Zheng, X.L., Zeng, D., Wang, F.Y.: Social balance in signed networks. Inform Syst Front 17, 1077-1095 (2015)
17. Majchrzak, A., Wagner, C., Yates, D.: The Impact of Shaping on Knowledge Reuse for Organizational Improvement with Wikis. Mis Quart 37, 455-470 (2013)
18. Beck, R., Rai, A., Fischbach, K., Keil, M.: Untangling Knowledge Creation and Knowledge Integration in Enterprise Wikis. Journal of Business Economics 85, 389-420 (2015)
19. Arazy, O., Ortega, F., Nov, O., Yeo, L., Balila, A.: Functional Roles and Career Paths in Wikipedia. Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, pp. 1092-1105. ACM, Vancouver, BC, Canada (2015)
20. Alavi, M., Leidner, D.E.: Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. Mis Quart 25, 107-136 (2001)
21. Jarvenpaa, S.L., Staples, D.S.: The Use of Collaborative Electronic Media for Information Sharing: an Exploratory Study of Determinants. The Journal of Strategic Information Systems 9, 129-154 (2000)
22. Wasko, M.M., Faraj, S., Teigland, R.: Collective Action and Knowledge Contribution in Electronic Networks of Practice. J Assoc Inf Syst 5, 2 (2004)
23. Johnson, S.L., Safadi, H., Faraj, S.: The Emergence of Online Community Leadership. Inform Syst Res 26, 165-187 (2015)
24. Gouldner, A.W.: The Norm of Reciprocity: A Preliminary Statement. Am Sociol Rev 25, 161-178 (1960)
25. Nowak, M.A., Sigmund, K.: Evolution of Indirect Reciprocity. Nature 437, 1291-1298 (2005)
26. Faraj, S., Johnson, S.L.: Network Exchange Patterns in Online Communities. Organ Sci 22, 1464-1480 (2011)
27. Block, P.: Reciprocity, Transitivity, and the Mysterious Three-Cycle. Social Networks 40, 163-173 (2015)
28. Rambaran, J.A., Dijkstra, J.K., Munniksma, A., Cillessen, A.H.N.: The Development of Adolescents' Friendships and Antipathies: A Longitudinal Multivariate Network Test of Balance Theory. Social Networks 43, 162-176 (2015)

29. Labianca, G.: Negative Ties in Organizational Networks. Contemporary Perspectives on Organizational Social Networks, pp. 239-259 (2014)
30. Contractor, N.S., Wasserman, S., Faust, K.: Testing Multitheoretical, Multilevel Hypotheses About Organizational Networks: An Analytic Framework and Empirical Example. Academy of Management Review 31, 681-703 (2006)
31. Blau, P.M.: Exchange and Power in Social life. Transaction Publishers (1964)
32. Emerson, R.M.: Social Exchange Theory. Annual Review of Sociology 2, 335-362 (1976)
33. Mercken, L., Snijders, T.A.B., Steglich, C., Vartiainen, E., de Vries, H.: Dynamics of Adolescent Friendship Networks and Smoking Behavior. Social Networks 32, 72-81 (2010)
34. Lu, Y.D., Jerath, K., Singh, P.V.: The Emergence of Opinion Leaders in a Networked Online Community: A Dyadic Model with Time Dynamics and a Heuristic for Fast Estimation. Manage Sci 59, 1783-1799 (2013)
35. Chen, Y.-R., Chen, X.-P., Portnoy, R.: To Whom do Positive Norm and Negative Norm of Reciprocity Apply? Effects of Inequitable Offer, Relationship, and Relational-Self Orientation. Journal of Experimental Social Psychology 45, 24-34 (2009)
36. Diekmann, A., Jann, B., Przepiorka, W., Wehrli, S.: Reputation Formation and the Evolution of Cooperation in Anonymous Online Markets. Am Sociol Rev 79, 65-85 (2014)
37. Heider, F.: Attitudes and Cognitive Organization. The Journal of Psychology 21, 107-112 (1946)
38. Cartwright, D., Harary, F.: Structural Balance - a Generalization of Heider Theory. Psychol Rev 63, 277-293 (1956)
39. Leskovec, J., Huttenlocher, D., Kleinberg, J.: Signed Networks in Social Media. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1361-1370. ACM (2010)
40. Snijders, T.A.B.: Stochastic Actor-Oriented Models for Network Change. J Math Sociol 21, 149-172 (1996)
41. Snijders, T.A.B., Lomi, A., Torlo, V.J.: A Model for the Multiplex Dynamics of Two-Mode and One-Mode Networks, with an Application to Employment Preference, Friendship, and Advice. Social Networks 35, 265-276 (2013)
42. Snijders, T.A.B.: The Statistical Evaluation of Social Network Dynamics. Sociological Methodology 31, 361-395 (2001)
43. Putzke, J., Fischbach, K., Schoder, D., Gloor, P.A.: The Evolution of Interaction Networks in Massively Multiplayer Online Games. J Assoc Inf Syst 11, 69-94 (2010)
44. http://R-Forge.R-project.org/projects/rsiena/ (Aufgerufen: 03.04.2016)
45. Robbins, H., Monro, S.: A Stochastic Approximation Method. Ann Math Stat 22, 400-407 (1951)
46. Wang, C., Butts, C.T., Hipp, J.R., Jose, R., Lakon, C.M.: Multiple Imputation for Missing Edge Data: A Predictive Evaluation Method with Application to Add Health. Social Networks 45, 89-98 (2016)
47. Steglich, C., Sinclair, P., Holliday, J., Moore, L.: Actor-based Analysis of Peer Influence in A Stop Smoking In Schools Trial (ASSIST). Social Networks 34, 359-369 (2012)
48. Schweinberger, M.: Statistical Modelling of Network Panel Data: Goodness of Fit. Brit J Math Stat Psy 65, 263-281 (2012)
49. Putzke, J., Fischbach, K., Schoder, D., Gloor, P.A.: The Coevolution of Network Structure and Perceived Ease of Use. In: Wirtschaftsinformatik Proceedings, pp. 96. (2013)

The Role of Information Elaboration for Co-Construction of Meaning during Idea Convergence: A Causal Mediation Analysis

Isabella Seeber

University of Innsbruck, Department of Information Systems, Innsbruck, Austria
isabella.seeber@uibk.ac.at

Abstract. Teams need to co-construct meaning to establish shared understanding about concepts when converging on the best ideas generated from crowdsourcing events. Facilitation interventions can aid the co-construction of meaning. The causal mechanism is believed to be the extent exchanged information is elaborated on. However, this mediating role has not been empirically confirmed in past research. Information elaboration in teams with and without facilitation intervention was tested with causal mediation analysis by drawing on data collected in a laboratory experiment. The findings suggest that facilitated teams had better information elaboration and co-construction than non-facilitated teams. Moreover, information elaboration could be identified as a strong causal mechanism through which facilitation interventions affect the co-construction of meaning. The study contributes to unravelling the black box of team processes through which this causal effect of facilitation intervention arises and helps fostering the design of improved automated feedback mechanisms.

Keywords: causal mediation analysis; co-construction; facilitation intervention; feedback; idea convergence

1 Introduction

In practice this means that teams need to identify 10 – 40 ideas out of hundreds and thousands of ideas [1-3]. It is difficult to deduce the value of an idea [2, 4], which makes idea convergence a demanding decision-making process [3]. Unlike traditional convergence teams, teams working with crowdsourced ideas need to converge on ideas that are not their own. This requires convergence teams to elaborate on raw idea descriptions without knowing how the idea came about [5]. For this purpose, they need to establish shared knowledge [6]. Facilitation techniques for idea convergence allow intervening into information elaboration processes in order to drive shared understanding [7]. Empirical evidence investigating *how* such facilitation interventions can achieve better convergence outcomes is scarce [6] particularly in crowd settings [5]. It is unclear to what extent facilitation interventions affect information elaboration among team members and if the extent of information elaboration is the *causal mechanism* for the co-construction of meaning, an antecedent of shared understanding.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Seeber, I. (2017): The Role of Information Elaboration for Co-Construction of Meaning during Idea Convergence: A Causal Mediation Analysis, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 697-700

This study contributes to closing this gap by investigating the mediating role of information elaboration in teams that converge on crowdsourced ideas with or without facilitation intervention.

2 Background and Hypotheses Development

The team learning processes of individual construction, co-construction, and constructive conflict have been established as antecedents of shared understanding [8]. The team learning process of co-construction is particularly relevant to idea convergence, which strives to establish a shared understanding of the meaning of concepts [9]. Over the last decade, facilitation techniques have been designed that directly aim at helping team members to converge on ideas [9]. By actively intervening into the execution of the process and the discussion, the facilitator affects information processing in teams [10], which, in turn, alters how the team co-constructs the meaning of shared concepts [6]. Thus, *H1: Teams with facilitation intervention will have better co-construction of knowledge than teams without facilitation intervention*.

Team members need to elaborate on exchanged information [11] in order to synthesize individual understandings into shared meaning [12]. While information elaboration is conceptualized in this study as exchanges about the nature of knowledge in different domains, such as viewpoints or beliefs, co-construction of meaning is about the joint construction from previous exchanges [13]. Facilitation stimulates such kinds of deeper information processing of team members with feedback cues. Thus, *H2: Teams with facilitation intervention will have more information elaboration than teams without facilitation intervention.*

It is not clear to what extent the team process of information elaboration is affected by facilitation intervention and in case it is affected, if the extent of information elaboration is the *causal mechanisms* that defines the co-construction of meaning, an antecedent of shared understanding [8]. It is argued that facilitators can intervene into interaction processes to avoid shallow processing of exchanged information [14] and keep the team's interaction on topic [10]. This can be accomplished by asking questions, clarifying statements, and co-creating artefacts representing their common understanding [15]. Thus, *H3: The effect of facilitation intervention on co-construction of meaning will be mediated by the extent of information elaboration in teams.*

3 Methods

A laboratory experiment was conducted to test the hypotheses. The treatment condition was instantiated with a facilitation technique from the design pattern language for collaborative work practices called thinkLets [16]. The external facilitators were trained by a professional facilitator and worked at the department as PhD-students or post-doctoral students. For their interventions, facilitators relied on a pre-tested and predefined script that included step-wise instructions how to run the convergence process and 26 prompts. Subjects were recruited from an undergraduate information systems course and were randomly assigned to the experimental condition. The task described a flooding event in the fictitious city called Norvos, which was based on another emergency response task [17]. All teams were supported by the collaboration software ThinkTank by GroupsSystems. Co-Construction of meaning was adopted from [8]. Information elaboration was measured with two items adopted and adapted from [18]. Two additional items were added. The control variables were collaborative orientation, working history, gender, and team size. Validity tests were satisfactory.

4 Results, Contribution, and Limitations

Hypothesis 1 ($F(1,84)=10.272$, $p < 0.05$) and 2 ($F(1,84)=11.454$, $p < 0.05$) were accepted and suggest that teams with facilitation intervention will differ from teams without facilitation intervention. H3 suggested that the extent of information elaboration is the causal mechanism through which facilitated teams will show higher co-creation of meaning. The result shows that the indirect effect (ACME) [19] due to information elaboration is in fact significantly mediating the relationship between the treatment and the outcome with an estimate of 0.439 ($p < 0.01$). Sensitivity analyses were conducted and also found that the causal mediation results seem to be robust to unmeasured confounders.This study found support for the mediating role of information elaboration. Findings of this study help to unravel the black box of team processes through which the causal effect induced by facilitation or feedback arises [20]. Given the increasing demand to design effective automated feedback mechanisms [10] into collaboration environments, it is important to understand what team processes are affected, how they change and if they change into what direction team processes should change. There are some limitations to consider that provide additional avenues for future work. First, the causal mediation analysis did not consider any moderating influence on the mediation path. Second, the construct information elaboration is a mix of items deduced from past research and self-developed items. Third, this study focused on a single causal mechanism, information elaboration.

5 Acknowledgements

The research leading to the presented results was partially funded by the Austrian Science Foundation (FWF Project Nr. J3735-G27).

References

1 von Krogh, G., and Raisch, S.: ‘Focus Intensely on a Few Great Innovation Ideas’, Harvard Business Review, 2009, 87, (10), pp. 32-32

2 Jouret, G.: ‘Inside cisco’s search for the next big idea’, Harvard Business Review, 2009, September

3 Bjelland, O.M., and Wood, R.C.: ‘An inside view of IBM's' Innovation Jam'’, MIT Sloan management review, 2008, 50, (1), pp. 32-40

4 Kornish, L.J., and Ulrich, K.T.: 'Assessing the Quality of Selection Processes', Available at SSRN 2499049, 2014

5 Kohn, N.W., Paulus, P.B., and Choi, Y.: 'Building on the ideas of others: An examination of the idea combination process', Journal of Experimental Social Psychology, 2011, 47, (3), pp. 554-561

6 Bittner, E.A.C., and Leimeister, J.M.: 'Creating Shared Understanding in heterogeneous work groups: Why it matters and how to achieve it', J. Manage. Inf. Syst., 2014, 31, (1), pp. 111-143

7 Davis, A., de Vreede, G.J., and Briggs, R.O.: 'Designing thinkLets for convergence'. Proc. AMCIS 2007 Proceedings, Paper 3582007 pp. Pages

8 Van den Bossche, P., Gijselaers, W., Segers, M., Woltjer, G., and Kirschner, P.: 'Team learning: building shared mental models', Instructional Science, 2011, 39, (3), pp. 283-301

9 De Vreede, G.J., Briggs, R.O., and Massey, A.P.: 'Collaboration engineering: Foundations and opportunities: Editorial to the special issue on the journal of the association of information systems', Journal of the Association for Information Systems, 2009, 10, (3), pp. 121-137

10 Wheeler, B.C., and Valacich, J.S.: 'Facilitation, GSS, and training as sources of process restrictiveness and guidance for structured group decision making: An empirical assessment', ISR, 1996, 7, (4), pp. 429-450

11 Peñarroja, V., Orengo, V., Zornoza, A., Sánchez, J., and Ripoll, P.: 'How team feedback and team trust influence information processing and learning in virtual teams: A moderated mediation model', Computers in Human Behavior, 2015, 48, pp. 9-16

12 Kozlowski, S.W., and Chao, G.T.: 'The dynamics of emergence: Cognition and cohesion in work teams', Managerial and Decision Economics, 2012, 33, (5-6), pp. 335-354

13 Baker, M.: 'A model for negotiation in teaching-learning dialogues', Journal of Interactive Learning Research, 1994, 5, (2), pp. 199

14 Rietzschel, E.F., Nijstad, B.A., and Stroebe, W.: 'The selection of creative ideas after individual idea generation: Choosing between creativity and impact', British Journal of Psychology, 2010, 101, (1), pp. 47-68

15 Singh, G., Hawkins, L., and Whymark, G.: 'An integrated model of collaborative knowledge building', Interdisciplinary Journal of E-Learning and Learning Objects, 2007, 3, (1), pp. 85-105

16 De Vreede, G.J., Kolfschoten, G.L., and Briggs, R.O.: 'ThinkLets: a collaboration engineering pattern language', International Journal of Computer Applications in Technology, 2006, 25, (2-3), pp. 140-154

17 Santanen, E.L., Briggs, R.O., and De Vreede, G.J.: 'Causal Relationships in Creative Problem Solving: Comparing Facilitation Interventions for Ideation', J. Manage. Inf. Syst., 2004, 20, (4), pp. 167-197

18 Limayem, M., and DeSanctis, G.: 'Providing Decisional Guidance for Multicriteria Decision Making in Groups', ISR, 2000, 11, (4), pp. 386-401

19 Keele, L., Tingley, D., and Yamamoto, T.: 'Identifying mechanisms behind policy interventions via causal mediation analysis', Journal of Policy Analysis and Management, 2015, 34, (4), pp. 937-963

20 Imai, K., and Yamamoto, T.: 'Identification and sensitivity analysis for multiple causal mechanisms: Revisiting evidence from framing experiments', Political Analysis, 2013, 21, (2), pp. 141-171

Inductive Reference Modelling Based on Simulated Social Collaboration

Andreas Sonntag[1], Peter Fettke[1], Peter Loos[1]

[1] German Research Center for Artificial Intelligence at Saarland University, Institute for Information Systems, Saarbrücken, Germany
{andreas.sonntag,peter.fettke,peter.loos}@iwi.dfki.de

Abstract. Organizations nowadays possess huge repositories of process models. Inductive reference modelling can save costs and time by reusing process parts of process models belonging to a common domain. The inductive development of a reference model for a large corpus of process models is a difficult problem. Quite a few, primarily heuristic approaches have been proposed to the research community that require an approximate matching between the single processes. With our approach, we introduce a new concept that brings in for the first time an abstract efficiency simulation of the social collaboration around knowledge-based process models. A reference model is assembled featuring at least the topological minimum requirements to be significantly more efficient than the input process models. Our evaluation indicates that the approach is able to generate reference process models that are more efficient than the input process models and at least as a reference model designed by an expert.

Keywords: Social Collaboration, Knowledge Work, Social Network Analysis, Reference model mining

1 Introduction

1.1 Research Motivation

Business process models, as event-driven process chains (PMs), are representatives of processes in an enterprise or in an organization. Current processes may be analysed in order to be improved. Organizations maintain huge repositories of these PMs. Therefore, the PMs have to be analysed as it is necessary for optimizing and managing the repositories. Typically, business analysts have specialized knowledge of the processes and provide expertise in modelling reference process models (RMs). These RMs enable reusability, modularity and avoiding redundancy by the design of new models which saves costs and time. They represent a best practice in a domain and can be extended for individual requirements [1]. The inductive development of universally applicable and reusable RMs (reference modelling in the following) is of great importance in the field of business process management as it seeks to infer generic models or patterns of a domain [2]. A strategy of reference modelling is to

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Sonntag, A.; Fettke, P.; Loos, P. (2017): Inductive Reference Modelling Based on Simulated Social Collaboration, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 701-715

identify similarities between individual PMs and to derive an abstracted PM [3]. In contrast to deductive approaches, general theories and concepts are not applied for the development of RMs [4]. The utilization of RMs can save time and costs towards single PM variants [17] which makes their execution efficient. The efficiency of a RM hence should be considered when aiming at developing RMs.

1.2 Related Work

Prior respecters of reference modelling are: [25] mine RMs by analyzing desired behavior from event log data. [27] and [26] develop a RM that has a minimal graph-edit distance respectively a minimal cost of change to given process variants. [7] and [30] apply genetic algorithms. [32] apply factor analysis to find statistical commonalities in the process structures. [33] assemble a RM from hierarchical clustered process fragment. [28] seek to identify similar frequent substructures among PMs. [29] derive a RM by gathering similarities of given PMs not only by their structure but by their behavior profile. [31] iteratively create a RM based on the proximity of node pairs.

The named previous automatic reference modelling approaches develop a RM "bottom-up" by comparing the node labels or PM topology. This comparison requires critical assumptions such as a certain model design quality, consistent syntax and language. The understanding of the semantics or similarity of process elements is subject to the process matching problem whose underlying graph matching problem is principally NP-complete [16]. There are heuristic approaches for the problem that achieve an approximate solution which is not generally unique and is afflicted with a loss of information [5].

1.3 Research Problem and Approach

Social collaboration in organizations is a competitive advantage as it is a driver for efficiency, time and money saving potential, product quality and knowledge diffusion [15]. Reference modelling has to be seen under several facets simultaneously in regard to quality and usability [18]. We belief that social collaboration around PMs is an important facet for reference modelling, especially in the context of knowledge-based processes. In the field of inductive reference process model mining, a social perspective for matching the process flow is missing. Prior approaches are primarily based on label matchings and similar graph structure detection in the process models. The influence of the topology of social collaboration between the humans that work around the process (performers) is neglected. Our research question thus is, what is an efficient collaboration topology between performers for a corpus of process models.

Therefore, we want to conceptualize and implement in this work an approach for the inductive development of a RM from a corpus of knowledge-based business process models applying a social perspective for matching the process flow. The resulting RM shall consist of an efficient collaboration topology for the performers working at their associated process functions. Following the concept of [9], the RM is a PM derived from a performer network (PN) that represents at least the minimum requirements of the PM to be efficiently executed. The PN for the RM development is optimized to fit best to the efficiency of the input PMs. The definition of efficiency can be found in section 2.2. The RM development and interpretation does not need a label matching. The resulting RM consists the minimum process elements for being as efficient as the PMs from it was derived. It can be tailored to concrete organizations and processes. Thereby, risks for the execution potential of a PM can be identified and reduced; needed key performers and their influence can be tracked; the structure of successful co-workership and knowledge/information transfer can be disclosed. The organization becomes more transparent and costs as well as managing effort are reducible.

Our approach relies on social network analysis and evolution strategy [10]. Its empirical evaluation follows [8] and the prototypical implementation is based on design research [14]. The validation of our concept faces three scenarios in which our prototypical implementation develops RMs from three PM corpora of different domains and two languages. The developed RMs are compared to the given respective gold RMs of each scenario. Gold RMs are designed by domain experts.

The paper is structured as follows: After fundamental concepts and terms are introduced, our approach is described followed by the validation of our concept with the experimental design, the evaluation and the discussion of our results. Afterwards, a conclusion brings this work to a close.

2 Fundamentals

2.1 Social Network Analysis

A social network represents the relation between agents. Agents are constituted as nodes and the links between them as edges. The agents are atomic units and they can only communicate and collaborate with agents that are directly connected to them. Social network analysis (SNA) seeks to disclose social interactions, political power and co-workership by abstracting from social relationships [6]. A social network has *n* agents/nodes and *m* edges. People tend to form clusters which are groups of directly connected individuals [20]. Those edges are called local edges. Edges between clusters are called global. The extend of this tendency can be measured by the average clustering coefficient of a network, *CC*. *CC* is the number of the actual- in relation to the possible number of edges between an agent's neighbours, averaged over all agents. The number of an agent's neighbours is called degree. The mean degree of a network *md* is the average number of all agents' neighbours. The number of agents on the shortest path between two gents is called path length. The network density *dens* is defined as $\frac{2m}{n*(n-1)}$. It describes the relation between existing and possible edges in the network which makes it a measure for being sparse. Being sparse is an essential property of many social networks as they tend to form relatively small butt dense clusters with few global edges between them [19]. Among others, this makes such social networks a "small world", in which all individuals, e.g. independent from geographic distances, are separated by only a few edges [20]. Another important peculiarity of social networks is the appearance of hubs as [22] suggested in theory- and [21] in an empirical study as they confirmed the appearance of hubs in a real organization over time. Hubs are nodes that have a much higher degree than the most other nodes. Hubs constitute hierarchy which is a necessary empirical implication of social networks. The outcome of actual relationships strongly depends on structural network properties [23]. The identification of critical joints in the network such as the collaboration of working individuals that are geographically separated and that may belong to different functional- and hierarchical positions, is of strategic importance for collaboration [24].

2.2 Efficiency of Process Models

The concept of performer networks (PNs) and their efficiency is introduced by [9]. Almost every business process needs social collaboration to be efficiently executed [12]. The concept of PNs connects the aim of business process management to design, model and execute efficient processes with the potential of SNA. Performers are agents working on process functions with a set of capabilities in a PN. Capabilities for this study are simplified as a mapping between a process function and a number indicating the extent of being capable/efficient to work at this function.

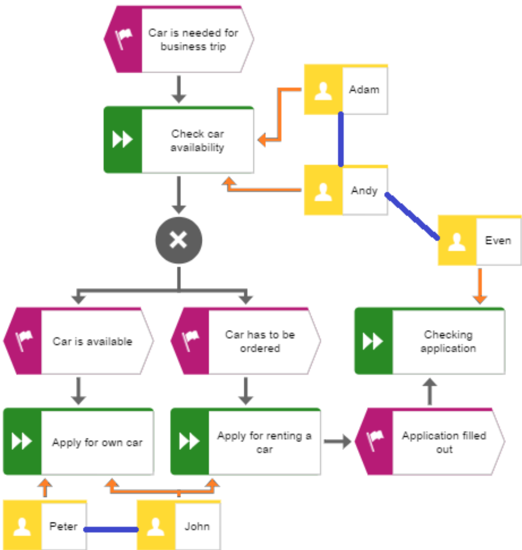


Figure 1. An example PN fragment around a PM

PNs represent a minimal topology of co-workership around business processes in an organization. PNs are formalized as social networks in which exist two kinds of edges: social edges that connect performers and functional edges that connect performers with process functions (*assigned_nodes*). An example PN fragment around a PM is presented in figure 1. The blue edges between the performers (named nodes) are social edges and the black edges/arrows between performers and process functions are functional edges. The other edges are process edges, connecting process nodes. The efficiency of a PM is the efficiency of the most efficient combination of a generated PN and the PM. The efficiency of a PN/PM combination is computed with an algorithm based on social network analysis and evolution strategy. Several repeated simulations of a work package, processed through the control flow of the PM, count how many iterations were needed to complete the whole process. Only performers associated to a process function (*involved_performers*) process the work package. They are much faster if they have the proper capability for this function. At the same time, neighbours of the performer collaborate which increases the PN efficiency PNE. PNE corresponds to the standardized PN efficiency definition of [9]. $PNE = 1 - (\frac{mean(\#neededIterations)}{max\#Iterations})$. PNE is the inverse of the relation between the number of needed iterations, averaged over all simulation runs $mean(\#neededIterations)$ and the maximal possible number of iterations. $max\#Iterations$ refers to a PN with one incompetent performer assigned to each process function. A PN can be generated by every social network generator that assigns at least one performer to a process function. The evaluation of PNs in [9] implies that a performer topology with hierarchy, short paths and significant clustering generates efficient PNs. Their suggested network generator by [11], PN generation algorithm and parameters are the basis for our PN generation in *generatePN* (see algorithm 5).

3 Approach

For our approach, we consider real-world PMs as event-driven process chains which are directed graph structures, consisting of a set of edges that indicates an order between a set of nodes. Each node has a label, as a linguistic expression of natural language, describing the node's function. Other (meta) information that might be provided is ignored. The fundamental assumption of our approach is that a RM must be efficient at least significantly more efficient than the most efficient input PM. We want to develop a reference PN from the input PMs entailing at least the topological minimum requirements to be efficient around all input PMs. It can be interpreted as a reference for positioning and putting people on team tasks for the whole organization in which all input PMs are assumed to be executed. Our approach's advantage is that no label matching or PM topology matching has to be considered as prior approaches require.

In this section we describe our approach on a high level. The procedure for the development of a RM from a set of input PMs is presented in algorithm 1. It is based on an evolutionary strategy which is an optimization driven only by mutating and selecting individuals [10]. The mutation rate can be increased every run, if no convergence was reached, which allows bigger steps through the solution space. The evolutionary strategy itself runs in linear time and can achieve a fitness/solution quality convergence to an acceptable solution after only a few iterations. The apparent advantage of evolutionary strategy, in contrast to other evolutionary algorithms, is to be fast and parallelizable because the population has only to be evaluated once per iteration. If it becomes foreseeable after many iterations that the individual's fitness will not improve, converges or even impairs, the evolutionary strategy can be run again or parallel runs on many processors at the same time can be executed until an acceptable solution occurs. The disadvantage on the other hand is that an evolutionary strategy can converge towards a local optimum in only few iterations but there is no proof that a global optimum will ever be reached [13].

Every PN is considered as an individual/possible solution. Its efficiency can only be seen in connection to a PM / or a set of PMs. The fittest resulting individual is the reference PN for the input PMs. In order to ensure reliability, the evolution strategy is repeated 1000 times as this meets the limitation of our computing time. We circumvent the label matching problem by mutating the individuals. The mutation is a random decision about how to arrange information inside an individual. This random decision made during the evolution compensates that we have no distinct knowledge about the distribution of the decision's quality. The selection forces the mutation quality to become better over many iterations. Therefore, a fitness function is defined in algorithm 4 that evaluates the sum of the efficiency of an individual towards all input PMs weighted by the inverse of its density. The efficiency of a PN/PM combination is evaluated with the procedure by [9]. The additional weight is necessary because otherwise a super PN with thousands of performers and functional edges would receive the highest fitness. A small/sparse PN with similar efficiency is to prefer. The same goes for a RM because the more process nodes a RM comprises, the bigger it becomes since there is no matching which would condense multiple nodes. As an objective evaluation criterion, a smaller RM with similar efficiency is likewise to prefer. "Super-models" with many nodes overloaded a user/modeller and prevented him from brining in own ideas [7].

The development of a RM follows the evolutionary strategy in *developReferenceModel* (algorithm 1): At first, for each PM in the set of input PMs *Models* an efficient PN is generated with *generatePN* in algorithm 5. Those PNs represent the initial population and are only optimized to be efficient for their respective PM. The best individual out of this population is picked by the selection operator who calculates all individuals' fitness values and chooses the individual with the highest fitness value. This individual is now called *best_individual* and its fitness is *best_fitness*. In the same step, *cur_fitness* and *minimum_fitness* are set to the same fitness value. The current individual *cur_individual* on which the mutation and selection operators are applied on is now created and initially set to a copy of *best_individual*. *cur_fitness* is the fitness of the current individual. The minimum fitness *minimum_fitness* is the fitness that must be significantly exceeded by *cur_fitness* during the evolution strategy (step 5) in which the mutation and selection operators drive *cur_individual* to reach a fitness convergence significantly greater than the minimum fitness. The mutation of an individual (algorithm 2) assigns random capabilities out of the possible capabilities in the input PMs to a randomly chosen portion of performers (between 0 and 100%) and flips their edges. When the evolution strategy found a good PN, a RM is derived from it (see *deriveEfficientModel* in algorithm 3): The RM contains all process functions to which the performers of the PN are assigned. A process edge between two functions is added if any pair of performers is connected so that the first performer is assigned to the first function and the other one to the second function. Conversely, that means that a generated PN around the RM is efficient in all combinations with the input PMs.

The developed RM is to understand as a minimum model that exhibits at least the efficiency of the input PMs and simultaneously it has relatively few nodes. By extending the fitness function, the resulting RM can be adapted to concrete requirements. The final step in the RM development would be to specify the performers in the real organization and bringing the reference PN and the developed RM together. The developed RM has to be seen as a combination of a reference PN with its corresponding PM. Only this combination can be interpreted and compared. This is a comparable constraint to a RM developed with the help of a label matching as that RM can only be used if the matching is of suitable quality which is subjective and difficult to measure.

Algorithm 1: Development of a RM:
developReferenceModel(Models)
Input: *Models*, A set of process models   Output: A RM

1. Let *best_individual* be a PN = *pn* that maximizes *fitness(pn, Models)* with *pn* = *generatePN(m)* for each PM *m* ∈ *Models*
2. *best_fitness* = *cur_fitness* = *minimum_fitness* = *fitness(best_individual, Models)*
3. *cur_individual* = *best_individual*
4. **Repeat** until convergence of *cur _fitness* >> *minimum_fitness*:
    a. *mutate(cur_individual, mutation_rate, Models)*
    b. *cur_fitness* = *fitness(cur_individual, Models)*
    c. **If** *cur_fitness* > *best_fitness* **Then**:
        i. *best_individual* = *cur_individual*
        ii. *best_fitness* = *fitness(best_individual, Models)*
5. **return** *deriveEfficientModel(best_individual)*

Algorithm 2: Mutation of a PN: *mutate(PN, mutation_rate, Models)*
Input: A PN, *PN*, *Models* and *mutation_rate* in [0;1]

1. *portion* = a set with (*mutation_rate* * 100%) of *PN.performers*
2. **For Each** *performer* in *portion* **Do**:
    a. *performer.capabilities* = random capabilities from *Models*
3. **For Each** edge *(p1, p2)* with *p1,p2* ∈ *portion* **Do**:
    a. flip edge *(p1, p2)*

Algorithm 3: Deriving an Efficient PM from a PN:
deriveEfficientModel(PN)
Input: A PN, *PN*   Output: A PM

1. *M* = empty PM
2. *M.add_nodes(PN.assigned_process.functions)*
3. **For Each** edge *(p1, p2)* **In** *PN.social_edges* **Do**:
    a. **For Each** (*n1, n2*) ∈ *(p1.assigned_nodes, p2.assigned_nodes)* **Do**:
        i. **If** there is at least one edge (*ip1*, *ip2*) in (*n1.involved_performers*, *n2.involved_performers*) **Then**:
            1. *M.add_edge(n1, n2)*
4. *M.delete_unconnected_nodes()*
5. **return** *M*

Algorithm 4: Fitness of a PN: *fitness(PN, Models)*
Input: A PN, *PN*   Output: A numeric fitness

1. **return** $\sum efficiency(PN, m) * (1 - dens(PN))$ for each *m* in *Models*

Algorithm 5: Efficient PN Generation around a PM Based on [9]: *generatePN(PM)*
Input: A PM, *PM*   Output: A PN

1. let *PN* be a random social network
2. assign *PN.performers* to random capabilities from *PM*
3. assign *PN.performers* to random process functions from *PM*
4. simulate *PN* around *PM*
5. **Repeat** 1-4 until *PN.efficiency* reaches a local optimum after an adequate number of iterations
6. **return** *PN*

4 Validation of Concept

4.1 Experimental Design

The validation of concept elucidates its potential and applicability. Herby, the concept approach is validated to be able to reproduce its results (reliability), to explain the increase of result quality (intern validity) and to produce a result that is generalizable/transferable (extern validity) [8]. For that purpose, we implemented our approach within the "Refmod-miner" framework which is a prototypical software platform (http://refmod-miner.dfki.de/). All algorithms described in the approach section were implemented in Java, the used network generators and SNA algorithms in section 2.1 were taken from the Python library NetworkX 1.11. The hardware configuration on which the implementation of our approach was executed involves an Intel(R) Core(TM) i7-3610QM CPU @ 2.30GHz and 8GB of RAM.

We provide three evaluation scenarios in which we demonstrate the reliability, intern validity and extern validity of our approach. The scenarios rely on three different real-world knowledge-based domains and two different languages, German and English. The three scenarios cover S1: 5 knowledge transfer PMs from knowledge management in outsourcing relationships and knowledge progress control [34], S2: 10 PMs created during the workshop for modelling in higher education (MoHoL 2016) [35]. S3: 7 PM variants/solutions about business trip admission for business informatics assignments in exams at a German university [36]. For each scenario, also a gold RM is provided, made by a domain expert, whose execution comes to an exemplary result comparing to the models in its scenario. For example, in S2, at the end, a customer is informed and in S3, a car for a business trip is ordered. PM properties, such as the sum, minimum value, first, second and third quartile and the maximum value for the number of nodes, start nodes, number of process edges and graph density, are presented in table 1.

Table 1. Overview over PM metrics: sum, minimum, maximum and quartiles

S1|**S2**|S3	sum	min	q25	median	q75	max
nodes	194|**187**|264	18|**15**|19	22|**17**|36	28.5|**18.5**|41	35|**19**|43	62|**26**|43
start nodes	6|**9**|11	1|**1**|1	1|**1**|1	1|**1**|2	1|**1**|2	1|**1**|2
edges	221|**200**|280	23|**17**|20	23|**18**|38	30|**20.5**|44	39|**21**|45	76|**25**|46
density in %	26|**62**|21	2|**4**|3	3|**6**|3	4|**6**|3	5|**7**|3	8|**8**|6

As described in the approach, our prototypical concept implementation runs *developReferenceModel(Models)* 1000 times. The mutation rate was set to 0.2 as this rate indicated significantly better results than 0.1, 0.3 and 0.4 in the first 10 runs. For each scenario corpus, each evaluation of PNs is repeated 100 times, and records these topological properties of the evolving individuals in every iteration: *num_per* (number of performers), *md* (average performer degree), *CC* (average clustering coefficient), *dens* (the PN density), *minimum_fitness* (defined in *developReferenceModel* as the fitness of the best PN for the input PMs, respectively the best initial individual (remains constant during the algorithm run), *cur_fitness* (fitness of the actual evolving individual), *best_fitness* (fitness of the present best individual). The fitness of the respective gold RM *gold_fitness* is measured to ensure that the gold RMs are comparable to our generated models. The fitness of a gold RM is evaluated by the same procedure as the minimum fitness is calculated, namely by the fitness sum of its most efficient PN towards all input PMs.

4.2 Evaluation

The runtime of *developReferenceModel* took on average 15 seconds per run. Runs in scenario S2 took on average 20% longer than the others and runs in S3 took on average 7% longer than S1. One iteration took between 0.25 and 2 seconds. The fitness of evolving individuals increases with the number of iterations in one run for reaching *cur_fitness (#iterations*) during the algorithm ($p < 0.001$) in all scenarios which can be observed in figure 2. Each point stands for a run of the PN development algorithm (the respective scenario that was run is marked with its own colour).

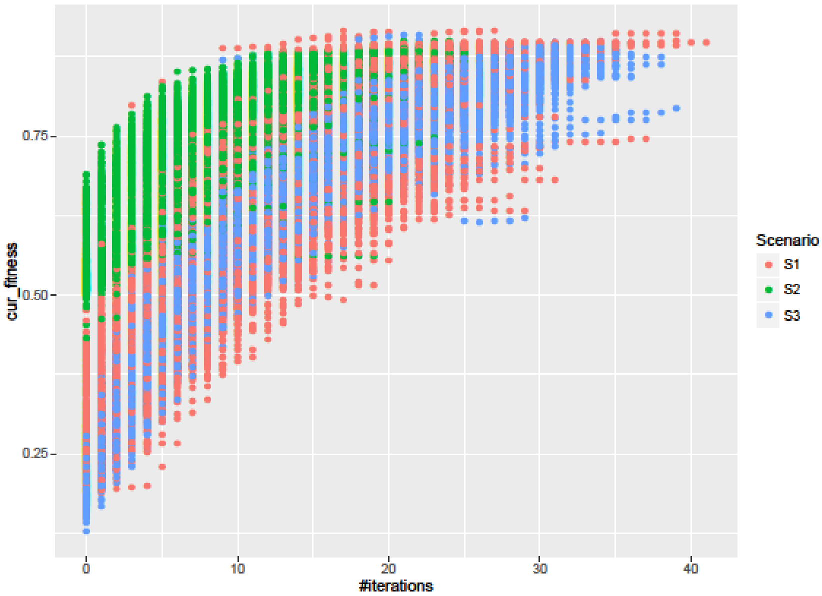


Figure 2. *#iterations* vs *cur_fitness* for all scenarios

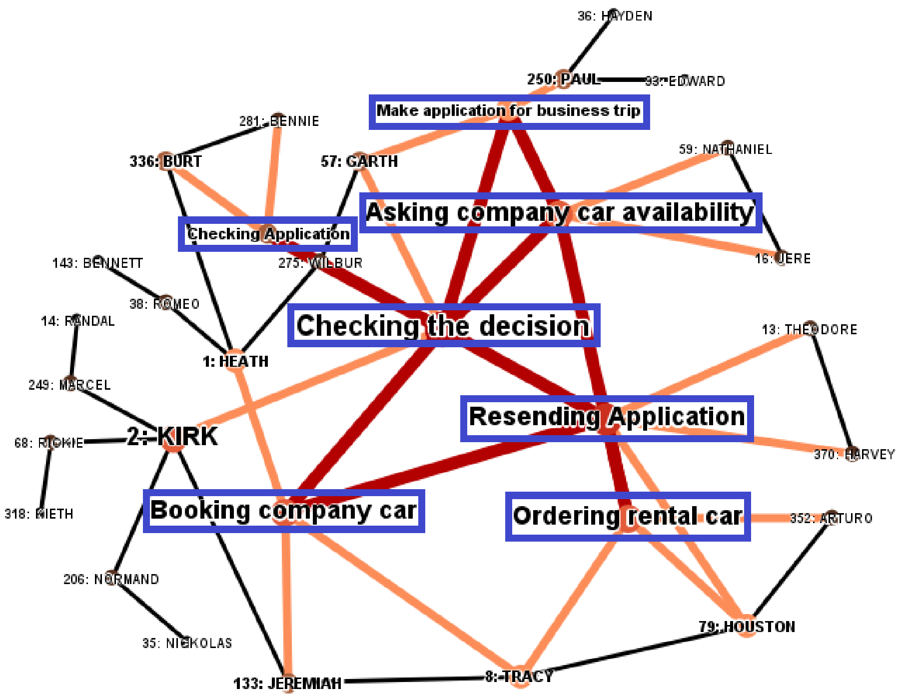


Figure 3. Reference PN with its derived RM for S3

Table 2. Descriptives for *best_fitness* over all runs with smallest value (Min), arithmetic mean (Mean), standard deviation (SD) and highest value (Max)

Scenario	Min	Mean	SD	Max
S1	0.84	0.89	0.01	0.92
S2	0.84	0.88	0.01	0.90
S3	0.84	0.88	0.01	0.91

Table 3. Average efficiency over all algorithm runs with standard deviation in brackets

Scenario	*best_fitness / minimum_fitness*	*best_fitness / gold_fitness*	*gold_fitness / minimum_fitness*
S1	4.67 (0.07)	1.19 (0.04)	3.92 (0.13)
S2	1.76 (0.02)	1.56 (0.04)	1.13 (0.03)
S3	6.01 (0.10)	2.58 (0.12)	2.33 (0.10)

Table 4. Correlations between respective variable and *cur_fitness*, $p < 0.01$

Scenario	*#iterations*	*num_per*	*md*	*CC*	*dens*
S1	0.86	0.86	-0.90	0.86	-0.88
S2	0.80	0.80	-0.82	0.77	-0.79
S3	0.92	0.93	-0.92	0.86	-0.88

In no run, more than 42 iterations were needed to reach convergence. As we can see in table 2, the mean best reached fitness lies nearer to its maximum than to its minimum. The standard deviation of *best_fitness* amounts to 0.01 for all scenarios. All PNs generated from the scenario respective gold RMs have a very low *num_per*, *CC* and *dens* with a simultaneously high efficiency and fitness over all input PMs. Figure 3 presents the fittest PN together with its derived PM, which we see as RM for all input PMs in S3. The figure contains process functions and process edges of the RM, performers of the reference PN and the related functional as well as social edges. All edges are undirected indicating mutual relations. Nodes with numbers are generated performers and the others with blue boxes are process functions. Black curves are social edges, red curves are process edges and orange curves are functional edges. Size and colour intensity of nodes and edges indicate the strength of their degrees. For reasons of clarity and comprehensibility only performers and process functions with a degree greater than 2 are plotted. In table 3, the average efficiency of all algorithm runs is presented. The average efficiency is the ratio between *best_fitness* and *minimum_fitness* of a run, averaged over all runs. The same applies for the relations between *best_fitness* and *gold_fitness* and *gold_fitness* and *minimum_fitness*. As an example, the fitness of the most efficient evolved PN in S3 was 2.58 higher than the respective gold PN and 6.01 times higher than the highest fitness of all individuals from the initial population (*minimum_fitness*). The fitness of the gold PN was 2.33 times greater than the minimum fitness. The relation between the fitness of the gold PN and the best initial individual was 2.33. Each standard deviation is significantly lower than a tenth of the corresponding average relation which speaks for a sufficient number of algorithm runs for this evaluation.

4.3 Discussion

The mutation rate of 0.2 reached a fitness convergence in all scenarios. Maximal 5 runs were needed to find an acceptable individual. The degree distributions of all efficient PNs around the gold RM and the best evolved PN, for all scenarios show the existence of hubs. Also, those PNs exhibit a significant but low clustering and density. The existence of hubs is an evidence for hierarchy which means that teams/clusters have single members that are much more central (between more people) than others [22]. This finding seems to be a condition for efficiency which is also confirmed by [9] for knowledge working processes. The number of performers correlates with the fitness over many iterations which means that more performers make the evolving PN more efficiently but this effect is compensated by the negative influence of the mean degree. All correlations in table 4 are stable over the scenarios in the meaning of their equal sign. The topological PN properties density, clustering and mean degree seem to be fitness drivers. For the approach's reliability pleads this stability after 1000 repetitions of the algorithm and of the repeated generation/evaluation of each PN.

According to our experimental design, we see the intern validity as confirmed by the significant fitness increase towards the minimum fitness for all scenarios over all runs (h0: *cur_fitness* < *minimum_fitness*; $p < 0.001$). Also there is a significant correlation between the number of iterations and the fitness ($p < 0.001$). Our algorithm so has an ascertainable influence on the evolvement of better individuals. For the extern validity, we tested the significant fitness increase towards the fitness of the gold reference PNs for all scenarios over all runs (h0: *cur_fitness* < gold_fitness; $p < 0.001$). The relation between the gold reference PN and the minimum fitness was always significantly greater than 1.0. That is an assertive indicator for our algorithm to be able to generate efficient reference PNs for various scenarios. All gold RMs were evaluated to be highly efficient and fit. This speaks for the validity of our approach because arbitrary models out of the initial population are significantly less efficient (h0: *fitness(generatePN(m), Models) > fitness(generatePN (gold_PM), Models)* for each PM *m* in the input PMs *Models* and the respective gold RM; $p < 0.01$). That means that the generated PN around the gold RM is efficient for all input PMs. Considering the low density and the random assignment of capabilities to performers in an efficient generated PN, its performers have only the most critical capabilities to work on the most critical process functions over all input PMs. This also means that the gold RMs' efficiency can be compared to other RMs for the input PMs which makes our fitness function a valid indicator for the quality of a RM.

The advantage of our approach is that the RMs can be generated valid over different domains. The proposed PN/RM combination describes a minimum topology of performers and their assigned process functions that is efficient towards the set of given input PMs. This proposed PN and RM can be adapted for specific stated requirements such as pre-given teams or pre-assigned performers to certain process functions. For a real environment, their efficiency comparing to other team/hierarchy constellations can then be simulated and evaluated. The RM in figure 3, as an example, indicates that the work flow between "Booking company car", "Ordering rental car" and "Checking the decision" is most critical for the efficiency of all process variants in S3 as they have the highest degrees of all process functions which makes them central in the RM. These process functions lie on critical paths in most process variants in S3, in the meaning of paths that reach from start nodes to end nodes and lie on many other paths at the same time. For that reason, most of all hub performers, such as "Kirk" "Jeremiah" and "Heath", were placed to work with their subordinated teams at this process functions. This can be interpreted as a recommendation or reference for a modeller to focus on needed capabilities for this process region when positioning real personal, e.g. at checking the decision for booking a car.

Limitations: Our approach focusing only on PN/PM topology is quite abstract and based on simplified assumptions about real processes and organizations. Organizations in our approach only consist of a set of performer networks and PMs. Their execution environment, social behaviour, resource allocation, communication- and production/processing capacities are not considered. In order to demonstrate the potential of our approach, the implemented algorithm produces quickly an acceptable result but will hardly reach a global optimum. For achieving a much better result, the number of performers should be reduced nearly to the number of needed performers in the gold RMs. An adapted, organization-specified implementation that considers the environment, the concrete performer capabilities and restrictions for their process assignment will be imperative for our approach to be utilized by practitioners.

5 Conclusion

In this paper we introduce a new concept for the inductive development of reference process models. A social perspective for matching the process flow is applied, rather than a traditional label matching which is an inexact and subjective approach. For a given set of input PMs, a reference process model is developed, in a few seconds of runtime, by including all process functions that are minimum requirements for the resulting model to be efficient. The efficiency is measured by the time that simulated performers need to complete the process. Three evaluation scenarios are provided to evaluate our approach. The evaluation indicates that the generated reference process models are at least as efficient as the input PMs and as a RM designed by an expert. Our results confirm the potential of our approach as they confirm its extern validity.

From a theoretical point of view that means that the efficiency of RMs designed by experts can be compared to our developed RMs which makes our fitness function a valid indicator for the quality of a RM. This in turn implies social collaboration to be an important facet for reference modelling. Our approach can be tailored to concrete organizations and processes. Practitioners take advantage of pre-selecting efficient sub corpora out of many models and identifying maybe invisible lead performers / critical junctures, in contrast to the formal structure, constituting efficient structures of co-worker ship around process models.

In future works, we want to provide a method to evaluate the quality of reference process models based on this approach. Also we will evaluate event logs of the execution of business processes to add a time/cost component to our fitness function for developing reference process models.

References

1. Becker, J. and Meise, V.: Strategy and Organizational Frame, Process Management. A Guide for the Design of Business Processes, J. Becker, M. Kugeler and M. Rosemann (eds.), Springer, 2011.
2. Fettke, P., Loos, P.: Perspectives on Reference Modeling. In: Fettke, P., Loos, P. (eds.) Reference Modeling for Business Systems Analysis, pp. 1-20. Idea Group, Hershey, PA., 2007.
3. Walter, J., Fettke, P., Loos, P.: How to Identify and Design Successful Business Process Models: An Inductive Method. In: Becker, J., Matzner, M. (eds.) Promoting Business Process Management Excellence in Russia - Proceedings and Report of the PropelleR 2012 Workshop, pp. 89-96. Moscow, Russia, 2013.
4. Becker, J., Schütte, R.: A Reference Model for Retail Enterprises. In: Fettke, P., Loos, P. (eds.) Reference Modeling for Business Systems Analysis, pp. 182-205. Idea Group, Hershey, PA, 2007.

5. Thaler, T., Hake, P., Fettke, P., Loos, P.: Evaluating the Evaluation of Process Matching Techniques. In: Kundisch, D., Suhl, L., Beckmann, L. (eds.) Tagungsband Multikonferenz Wirtschaftsinformatik 2014, MKWI-2014, pp. 1600-1612. Paderborn, Germany, 2014
6. Wasserman, S. and Faust, K., Social network analysis: Methods and applications, volume 8. Cambridge university press, 1994.
7. Martens, A., Fettke, P., Loos, P.: A Genetic Algorithm for the Inductive Derivation of Reference Models Using Minimal Graph-Edit Distance Applied to Real-World Business Process Data. In: Kundisch, D., Suhl, L., Beckmann, L. (eds.) Tagungsband Multikonferenz Wirtschaftsinformatik 2014, MKWI-2014, pp. 1613-1626. Paderborn, Germany, 2014.
8. Campbell, DT., and Stanley, JC.: Experimental and quasi-experimental designs for research. Ravenio Books, London, 2015.
9. Sonntag, A., Fettke, P.: Efficiency Of Generated Performer Networks In Collaborative Business Process Models. In: IEEE Conference on Business Informatics (CBI). Conference on Business Informatics (CBI-16), August 29 - September 1, Paris, France, IEEE, 2016.
10. Beyer, H-G.: The theory of evolution strategies, Springer Science & Business Media, 2013.
11. Holme, P. and Kim, B. J.: Growing scale-free networks with tunable clustering, Physical Review E, 65(2): 026107, 2002.
12. Niehaves, B. and Plattfaut, R.: Collaborative business process management: status quo and quo vadis, Business Process Management Journal, 17(3): 384–402, 2011.
13. Beyer, HG, and Schwefel, HP.: Evolution strategies–A comprehensive introduction. Natural computing 1.1: 3-52, 2002.
14. Hevner, A., March, S., Park, J. and Ram, S.: Design science in information systems research, MIS Quarterly, Vol. 28 No. 1, pp. 75–105, 2004.
15. Cross, R., Borgatti, S. P., and Parker, A.: Making invisible work visible: Using social network analysis to support strategic collaboration, California management review, 44(2):25–46, 2002.
16. Garey, M.R., Johnson, D.S.: Computers and Intractability: a Guide to the Theory of NP-Completeness, Freeman, San Francisco, 1979.
17. Scheer, AW, and Nüttgens, M.: ARIS architecture and reference models for business process management. Business Process Management. Springer Berlin Heidelberg, 376-389, 2000.
18. Fettke, P. and Loos, P.: Multiperspective evaluation of reference models–towards a framework. International Conference on Conceptual Modeling. Springer Berlin Heidelberg, 2003.
19. Barabási, AL.: Network Science. Cambridge University Press. Retrieved 25 May, 2015.
20. Watts, DJ.: Six degrees: The science of a connected age, WW Norton and Company, New York, 2004.
21. Leskovec, J., Kleinberg, J., and Faloutsos, C.: Graphs over time: densification laws, shrinking diameters and possible explanations, In Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, pages 177–187, ACM, 2005.
22. Barabási, A.-L. and Albert, R.: Emergence of scaling in random networks, science, 286(5439):509–512, 1999.
23. Johnson-Cramer, M. E., Parise, S., and Cross, R. L.: Managing change through networks and values, California Management Review, 49(3):85–109, 2007.

24. Cross, R., Borgatti, S. P., and Parker, A.: Making invisible work visible: Using social network analysis to support strategic collaboration, California management review, 44(2):25–46, 2002.
25. Gottschalk, F., Van Der Aalst, W., Jansen-Vullers, M.: Mining reference process models and their configurations. In: Meersman, R., Tari, Z., Herrero, P. (eds.) On the Move to Meaningful Internet Systems: OTM 2008 Workshops. Lecture Notes in Computer Science, vol. 5333, pp. 263–272. Springer, Berlin, 2008.
26. Ardalani, P., Houy, C., Fettke, P. and Loos, P.: Towards a Minimal Cost of Change Approach for Inductive Reference Model Development, Proceedings of the 21st European Conference on Information Systems, AIS, Utrecht, 2013.
27. Li, C., Reichert, M. and Wombacher, A.: Discovering Reference Models by Mining Process Variants Using a Heuristic Approach, in Dayal, U., Eder, J., Koehler, J. and Reijers, H. (Eds.), Business Process Management: 7th International Conference, BPM 2009, Ulm, Germany, September 8-10, 2009. Proceedings, Vol. 5701, Springer, Berlin, Heidelberg, pp. 344–362, 2009.
28. Rehse, JR., Fettke, P. and Loos, P.: A graph-theoretic method for the inductive development of reference process models, Software & Systems Modeling, 2015.
29. Rehse, JR.; Fettke, P.; Peter Loos, P.: An Execution-Semantic Approach to Inductive Reference Models Development, in: 24th European Conference for Information Systems (ECIS-16), June 12-15, Istanbul, Turkey, Association for Information Systems (AIS), 2016.
30. Yahya, B.N., Wu, J.-Z. and Bae, H.: Generation of Business Process Reference Model Considering Multiple Objectives, Industrial Engineeering & Management Systems, Vol. 11 No. 3, pp. 233–240, 2012.
31. Yahya, B.N. and Bae, H.: Generating Reference Business Process Model Using Heuristic Approach Based on Activity Proximity, Intelligent Decision Technologies, Springer, pp. 469–478, 2011.
32. Martens, A., Fettke, P. and Loos, P.: Inductive Development of Reference Models Based on Factor Analysis, in Thomas, O. and Teuteberg, F. (Eds.), Proceedings Der 12. Internationalen Tagung Wirtschaftsinformatik (WI 2015), Vol. 12, Universität Osnabrück, Osnabrück, Osnabrück, Germany, pp. 438 – 452, 2015.
33. Ling, J. and Zhang, L.: Generating Hierarchical Reference Process Model Using Fragments Clustering, Asia-Pacific Software Engineering Conference (APSEC). IEEE, 2015.
34. Novara, C., and Schwabe, G.: Wissensmanagement in Outsourcingbeziehungen und Wissenskulturfortschrittskontrolle, Chur Schweiz, 2006.
35. Workshop for modeling in higher education (MoHoL 2016), http://butler.aifb.kit.edu/MoHoL/?page_id=62 (Accessed: 21.10.2016)
36. Repository of business informatics exams solutions, http://rmm.dfki.de/index.php?site=repository&file=Exams&source=repo (Accessed: 21.10.2016)

Functional Interdependencies between Quality Techniques reverting to Meta Models

Florian Johannsen[1]

[1] University of Regensburg, Management Information Systems, Regensburg, Germany
florian.johannsen@wiwi.uni-regensburg.de

Abstract. Considering the digitalization of the economy, process-oriented quality management (PQM) has increasingly been gaining attention. In the course of a PQM project, quality techniques are applied to elicit employees' process knowledge and transform it into solutions to overcome process weaknesses. However, quality techniques may support each other during application or produce contradictory results, depending on the so-called "functional interdependencies (FIs)" between them. Little understanding exists of how such FIs can be properly identified, which is a prerequisite to exploit valuable synergies between quality techniques. To uncover the corresponding interdependencies, we revert to meta models in this paper, which allow to precisely describe a technique's functionality. Generally valid indicators on a meta model level are derived to unveil the existence of FIs.

Keywords: Process knowledge, meta model, quality technique.

1 Introduction

Many enterprises go through profound transformations these days triggered by the increasing digitalization of the economy [1]. Against this background, the improvement or redesign of business processes, in the context of process-oriented quality management (PQM) projects, is an important task [2], [3]. Only if the business processes are aligned with the expectations of internal and external customers, the purposeful definition of business services and the introduction of IT systems to beneficially support a company's value creation are possible [4].

However, the execution of PQM projects is challenging and many initiatives fall short of expectations [2], [5]. The success of PQM projects largely depends on the participation of employees from all cooperating partners in an inter-organizational business network and the goal-oriented elicitation [6] of their process knowledge to derive opportunities for process improvement (cf. [7]).

In this respect, the PQM discipline has brought forth a variety of methods (e.g., Six Sigma) that can be applied to improve or redesign business processes [3]. However, many employees do not have the time to become acquainted with such holistic methods (cf. [8]). Further, their application is increasingly perceived as too resource-consuming for projects with a limited scope (cf. [9]). Thus, enterprises often prefer a manageable

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Johannsen, F. (2017): Functional Interdependencies between Quality Techniques reverting to Meta Models, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 716-730

and limited set of quality techniques (e.g., Ishikawa Diagram, etc.) instead of extensive quality management methods (cf. [9]), even though the selection of adequate quality techniques is time-consuming (cf. [10]). Further, there are "functional interdependencies (FIs)" between quality techniques, i.e., they may complement each other during application or pursue diverging goals, e.g., cost-orientation vs. customer-orientation [11]. The knowledge of these interdependencies is decisive to understand which quality techniques may be purposefully combined in a project. For example, the CTQ/CTB-Matrix (Critical-to-Quality/Critical-to-Business-Matrix) helps to define quality goals, which are the base for developing Key Performance Indicators (KPIs) to measure process performance, e.g., via the Measurement-Matrix (cf. [12]). However, little understanding exists on such fertile FIs between quality techniques in practice and literature (e.g., [8], [11]). This is a drawback because such knowledge is crucial for method engineers to develop enterprise-adapted PQM methods and the corresponding software support (e.g., [13]). Therefore, our research aims at finding indicators that point to valuable FIs between quality techniques. For that purpose, we revert to meta models (cf. [14]). Meta models are suitable to capture the core concepts of quality techniques and to explicate their functionalities to transform particular types of input information (e.g., customer requirements) to output information (e.g., project goals) (cf. [10]). Accordingly, meta models allow to describe the nature of quality techniques and to derive indicators as to whether quality techniques can be beneficially combined in the course of a project or not. Thus, the following research question (RQ) is posed: *How can indicators of functional interdependencies (FIs) between quality techniques be purposefully identified on a meta model level and what FIs do typically exist between quality techniques of a representative set?*

Based on the findings, quality techniques with beneficial synergies can be precisely identified by reverting to meta models. Individual PQM approaches and corresponding software tools may then be designed.

This paper is structured as follows: first, foundations of quality techniques and meta models are explained. Afterwards, the research procedure is described and the design of meta models is explicated. In the main part of the paper, the indicators of FIs are derived. Afterwards, their benefits are presented. The paper concludes with an outlook.

2 Foundations

2.1 Quality Techniques and Functional Interdependencies (FIs)

A quality technique is a guideline for the creation of results in PQM projects (cf. [12], [15]). In that context, a quality technique supports the elicitation of employees' process knowledge (cf. [6]), derives some type of input information that is transformed to output information and thus (partial) results (cf. [10]). This perception is similar to the definition of a "technique" in IS method engineering (cf. [15]). An example for a quality technique is the Ishikawa Diagram (Fishbone Diagram), which serves the classification of problem causes for insufficient process performance (cf. [16]). According to *Bruhn* [11], functional, temporal as well as hierarchical interdependencies exist between quality techniques. Temporal interdependencies refer to the chronological sequencing of quality techniques in projects [11]. Hierarchical interdependencies differentiate whether a quality technique pursues strategic (e.g., definition of business objectives) or operational goals (e.g., definition of KPIs) [11]. Regarding the identification of valuable synergies between quality techniques, however, FIs are of interest (cf. [11]). FIs analyze the conjoint application of quality techniques considering their individual functioning, i.e., the way each technique converts input to output information, as well as the nature of the input and output itself. By taking this detailed perspective on how information is processed, it becomes obvious whether techniques may complement each other, produce identical outcomes or even require their mutual application (cf. [11]). Considering this, FIs are suitable to describe the interplay between quality techniques in the course of a project, and different types can be distinguished (Table 1).

Table 1. Types of FIs

Interdependency type (derived from [11])	Examples
Conditional: A quality technique requires other techniques to be applied in a project in addition.	Creating a Data Collection Plan requires the definition of KPIs in advance, e.g., by means of the Measurement-Matrix (e.g., [12]).
Complementary: Two or more quality techniques support each other during application.	The Failure-Mode-and-Effects-Analysis (FMEA) is purposefully complemented by the KANO Model, as the severity of potential "defects" during process execution can be precisely quantified that way (e.g., [12], [32]).
Substituting: The application of two or more quality techniques leads to identical types of output information.	The CTQ/CTB-Matrix as well as the Driver Tree can be used for specifying process-related quality goals (CTQ/CTB factors) for instance (e.g., [12], [17]).
Rivalling: The application of particular quality techniques produces results that are contradictory to one another.	Applying the FMEA and the Value-Stream-Map (VSM) may generate contradictory suggestions on the should-be process design (e.g., [12]).
Note: *Bruhn* [11] also mentions "indifferent interdependencies" indicating that the application of certain techniques has no mutual influence on one another. However, this type is not further considered because the corresponding techniques are not interrelated in terms of above described specification.	

Against this background, the beneficial synergies between quality techniques are primarily determined by conditional and complementary interdependencies. In case of a conditional interdependency, the application of a quality technique produces output information (e.g., KPIs) that is taken up and further processed as input information by another technique. In case of a complementary interdependency, the combined usage of certain techniques leads to results that are more precise in nature (e.g., KPIs aligned with project goals), while inherent drawbacks of a quality technique can be mitigated at the same time (see Table 1).

2.2 Meta Models and Process Knowledge in PQM Projects

Meta modeling is a widely-established discipline in the field of model-driven design and development of IS and software, respectively [14], [18]. In this regards, the behavior of software or an IT system is specified via conceptual models [18]. The meta model defines the concepts that can be expressed in such conceptual models designed with the help of a modeling language, i.e., what modeling elements may be applied [14]. Hence, a conceptual model is created as an instance of the corresponding meta model [13]. In the research at hand, we use meta models to describe the constituting elements (core concepts) of a quality technique as well as the type of input information (e.g., project goal) that is transformed to a particular type of output information accordingly (e.g., improvement idea). In so doing, an unambiguous description of a technique's functionality is achieved. Applying quality techniques in the course of projects results in diagrams, tables or sketches (cf. [19]), allowing to capture emerging process knowledge as conceptual models, which are specified by meta models accordingly. Thus, the documented result (e.g., conceptual model) received by applying a particular technique represents an instance of the corresponding meta model. For the design of meta models, we generally revert to UML class diagrams, which have proven suitable for creating meta models in research and practice alike (e.g., [20]).

Process knowledge plays a decisive role in light of organizational learning (cf. [7], [21], [22]). In this regards, "learning" specifies a firm's efforts to "retain and improve competitiveness, productivity, and innovativeness in uncertain technological and market circumstances" [22, p. 378]. Process knowledge is a multi-dimensional construct, comprising knowledge about the process structure, the training and management efforts required for achieving desired outcomes or knowledge directly linked to process execution (e.g., factors impacting efficiency) amongst others [21]. The challenge in PQM projects is to elicit employees' tacit process knowledge and convert it into explicit knowledge (cf. [23]) that can be used to derive opportunities for process improvement. Quality techniques support this conversion because employees' ideas, which are based on their individual process knowledge, are used as input to be transformed into results visualized as diagrams or conceptual models for instance (e.g., [19]). Meta models are suitable to capture the core concepts of quality techniques and to explain how the aforementioned conversion is performed. Further, the combination of discrete pieces of explicit knowledge [23] to come to improvement suggestions is fostered by techniques since the information processed may stem from diverse sources such as reports or IT-systems (e.g., [12]). The research contributes to developing means to uncover FIs between quality techniques and thus to support the purposive externalization and use of process knowledge to improve process performance.

3 Procedure of the Research

The study at hand is part of a larger Design Science (DS) project (cf. [24]), which aims at the development of a modeling tool to document, communicate and analyze knowledge on process weaknesses and process improvement opportunities. A central requirement on the tool is to support users in the selection and combined application of quality techniques based on FIs. For the implementation, the meta modeling platform ADOxx *(www.adoxx.org)* will be reverted to, which has been successfully applied in industry for more than 15 years now [13]. However, prior to the implementation, indicators for FIs on a meta model level are to be identified that allow to unambiguously decide whether quality techniques complement one another or not. Our paper deals with the identification of corresponding indicators and follows a four-step procedure building on the principles of inductive logic (cf. [25]). Thus, based on a sample set of quality techniques, the corresponding meta models are analyzed to derive generally valid indicators explaining the occurrence of FIs (see Fig. 1).

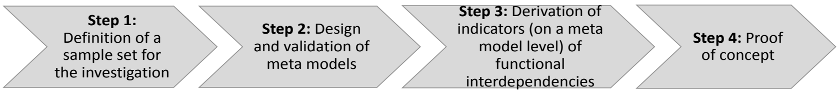


Figure 1. Procedure for identifying indicators

Considering the huge number of existing quality techniques (e.g., [12]), a representative set (sample set) is selected as a subject of investigation in a first step **(step 1)**. The techniques of our set along with the knowledge of the interdependencies between them represented the "instances" of the "phenomenon" [25] investigated. Afterwards, meta models for the quality techniques are generated and validated (**step 2**). Indicators explaining the occurrence of FIs between techniques are derived in **step 3**. Their applicability is demonstrated in **step 4** reverting to a prototypical realization.

4 Sample Set, Design and Validation of the Meta Models

4.1 Sample Set of the Investigation (Step 1)

On the one hand, the toolbox of quality techniques of a German automotive bank was reverted to for this research, comprising 30 techniques in total. This bank has a long tradition regarding the adaption and usage of methods for PQM (e.g., Lean Six Sigma, Total Quality Management, GE Work-Out) making it a suitable candidate for the investigation. Further, the author of this study participated in various PQM projects at the bank, taking the role of a "team member", over a period of three years. That way, profound insights into the beneficial combination of quality techniques were gained.

On the other hand, publications explicating FIs between quality techniques, e.g., in the form of an efficient further processing of results (conditional interdependency), were reverted to that were derived from a previously conducted literature review on the integration of quality methods and techniques (cf. [26]). Considering these findings, techniques not considered by the toolbox of the automotive bank were added to our sample set for the study, comprising 34 techniques in the end (see Appendix).

Next, the FIs between these quality techniques were specified according to the types as introduced in Table 1. That way, a complete overview of the FIs for the sample set emerged. For that purpose, the descriptions in literature were reverted to as well as the insights gained by actively participating in projects at the automotive bank. The Appendix exemplifies the results of this process for the CTQ/CTB-Matrix or the Driver Tree (cf. [17]) amongst others. The derivation of FIs was performed by two researchers, who consolidated the results afterwards. Further, the findings on FIs were validated in discussions with leaders of PQM initiatives at the automotive bank.

4.2 Design and Validation of the Meta Models (Step 2)

The subsequent design of meta models for the quality techniques of the set was done as follows: first, the core concepts of a quality technique were identified by analyzing its functioning. For example, the CTQ/CTB-Matrix supports the user in defining "Critical-to-Quality (CTQ)" and "Critical-to-Business (CTB)" factors based on customer (Voice of the Customer – VOC) and employee requirements (Voice of the Business – VOB), which are classified into core statements correspondingly (cf. [12]). Then, we considered the relations between the core concepts. For instance, each VOC or VOB statement in a CTQ/CTB-Matrix is assigned to one core statement at least. The core concepts and the relations were transformed into corresponding classes and relations of a meta model (MM) afterwards. Finally, the cardinalities of the meta model were to be set. Fig. 2 shows the meta model for the CTQ/CTB-Matrix.

A decisive aspect concerns the validation of meta models. In this context, formalization is an established means of uncovering inconsistencies, syntactical errors and incompleteness of the meta model design [27]. A generally valid formalization approach for domain-independent meta models is FDMM (Formalism for Describing ADOxx Meta Models and Models) (cf. [28]). Due to its general applicability across domains, differentiating FDMM from formalization approaches such as EMOF or KM3, which were developed for specifying software architectures in particular, it was chosen for the study at hand. All meta models established for the quality techniques of our sample set were formalized via FDMM making it possible to check them for inconsistencies (e.g., wrong cardinalities), syntactical errors (e.g., in case the meta model of the CTQ/CTB-Matrix would allow to connect instances of the class "VOC statement" to instances of the class "CTQ factor") and incompleteness (e.g., missing cardinalities). Generally, meta models in FDMM are represented as a tuple of a set of object types (O_i^T), data types (D_i^T) and attributes (A_j) [28], which is exemplified for the CTQ/CTB-Matrix in equation (1) (see [28] for details on FDMM).

$$\mathrm{MT}_{CTQ/CTB-Matrix} = \langle O^T_{CTQ/CTB-Matrix}, D^T_{CTQ/CTB-Matrix}, A_{CTQ/CTB-Matrix} \rangle \quad (1)$$

Further, the meta models were discussed with four researchers renowned for their expertise in the field of meta modeling and two practitioners who had a consultancy industry background and had been heavily involved in PQM projects for several years. The discussion partners were not involved in the development of the meta models and thus unbiased. Their feedback was gathered and modifications were made, if necessary. After these steps, the validity of the meta models was seen as sufficiently confirmed.

5 Indicators of Functional Interdependencies (Step 3)

Based on the meta models created in the prior step *(step 2)* and the FIs between the techniques of our sample set *(step 1)*, indicators of FIs on a meta model level were derived. Therefore, for each type of FI (see Table 1), we specified the synergies between the corresponding quality techniques of our sample set more profoundly and, if different forms of synergies could be distinguished, we defined subtypes of a FI *(step 3.1)*. Afterwards, each subtype (or form of synergy) was analyzed in detail by reverting to the meta models of the quality techniques, which allowed for explaining the FI by means of the meta models' classes *(step 3.2)*. Based on these findings, indicators on a meta model level could be derived characterizing each type of FI *(step 3.3)*. To reduce complexity, the indicators analyze quality techniques following a binary strategy (cf. [29]), i.e., it is determined whether two particular quality techniques considered hold an interdependency or not. The three steps were performed by two researchers – to reduce subjectivity – with the results being discussed and consolidated afterwards.

5.1 Conditional Interdependencies

Conditional interdependencies exist in case two or more quality techniques presuppose their mutual application (cf. [11]).

Step 3.1: In our sample set, conditional interdependencies between techniques were characterized by one particular form of synergy. In this respect, a technique produced a certain type of output information, which represented a specific type of input information for another technique simultaneously. For instance, the CTQ/CTB-Matrix provides the "CTQ" and "CTB factors" as types of output information (cf. [12]). These are referred to by the Measurement-Matrix for the definition of KPIs (cf. [12]). The use of the Measurement-Matrix is thus bound to techniques enabling the derivation of CTQ and CTB factors, e.g., the CTQ/CTB-Matrix (a more comprehensive example can be found at: *http://tinyurl.com/zbctxt8*).

Step 3.2: On the level of meta models (MM), the aforementioned synergy between quality techniques is visually exemplified in Fig. 2 reverting to the CTQ/CTB-Matrix and the Measurement-Matrix. The dotted line highlights the common classes across the meta models. It becomes obvious that the classes "Critical-to-Quality (CTQ) factor" and "Critical-to-Business (CTB) factor" can be found in both meta models, while they represent output information types (colored "black") in MM_1 (meta model of the CTQ/CTB-Matrix) and input information types (colored "grey") in MM_2 (meta model of the Measurement-Matrix). In this regards, the CTQ or CTB factors captured by an instance of MM_1 (CTQ/CTB-Matrix) serve as input information for an instance of MM_2 (Measurement-Matrix). This kind of relation between classes was similarly observed for all other techniques of the sample set regarding conditional interdependencies.

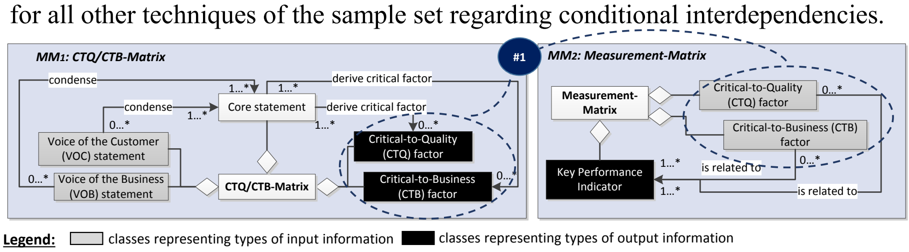


Figure 2. Example for conditional interdependencies (indicator #1)

Step 3.3: Based on these insights, the following indicator pointing to a conditional interdependency between two quality techniques, on a meta model level, was derived:

Indicator #1: A conditional interdependency between two *quality techniques i* and *j* is given, if the corresponding meta models (MM_i and MM_j) share identical classes while these represent types of output information of *quality technique i* in MM_i and types of input information of *quality technique j* in MM_j (or vice versa).[1] The results of the *technique i* thus represent decisive input information for *quality technique j* (or vice versa).

[1] In the following, *i* and *j* are continuous indices with the following assumptions: 1) $i \in \mathbb{N}$, 2) $j \in \mathbb{N}$ and 3) $i \neq j$. *Technique i* or *j* are thus representatives for any kind of quality technique.

5.2 Complementary Interdependencies

Generally, complementary interdependencies are observed for quality techniques that support each other during application (cf. [11]). Detailed examples of each subtype as described in the following can be found at: *http://tinyurl.com/zbctxt8*.

Step 3.1: In total, we identified four different subtypes (forms of synergy) of complementary interdependencies (A to D) between techniques based on our set.

Subtype A: The first subtype of complementary interdependencies builds on the use of common concepts, by different quality techniques, to transform input information to output information. For instance, both the KANO Model and the Driver Tree use the KANO categories "basic requirements", "breakthrough customer needs" and "core competitive requirements" [30], [31] to enable the prioritization of customer and employee requirements on process execution. The KANO Model prioritizes process-related VOC and VOB statements and its application thus clarifies which requirements are of particular interest to process stakeholders. Based on these insights, KPIs can be developed (Driver Tree) considering requirements with a high priority in special [17].

Subtype B: A second subtype of complementary interdependencies is given in case quality techniques contain core concepts that pursue a common purpose of transforming input information to output information (e.g., analysis, classification, comparison of information, etc.) [10], but are not identical per se. For example, the FMEA may be applied to identify potential "defects" during process execution and to judge their severity [12]. A classification of customer or employee requirements according to the KANO categories (cf. [30]) provides hints as to which potential "defects" will most probably affect customer satisfaction in a negative way [32]. Thus, a complementary interdependency between the KANO Model and the FMEA exists.

Subtype C: Further, a complementary interdependency exists, if the results gained by applying a particular quality technique help to specify the input information processed by another technique more precisely. However, the results produced by the first technique are not a mandatory prerequisite for applying the second one, which demarcates *subtype C* from a conditional interdependency. In a project for instance, the results generated via the CTQ/CTB-Matrix (CTQ and CTB factors) may trigger the purposeful search for causes of insufficient process performance reverting to the Ishikawa Diagram (cf. [16]). However, the process weaknesses to be investigated by means of the Ishikawa Diagram do not necessarily have to be derived from the CTQ or CTB factors but can also be defined "ad-hoc" in the course of a project (cf. [16]).

Subtype D: Yet another subtype addresses the usage of a quality technique to further refine the output information generated by another technique. An example would be the combined use of the Affinity Diagram and the Payoff-Matrix (cf. [12], [33]). By using the Affinity Diagram, suggestions for process improvement are purposefully classified (e.g., cost-oriented solutions, IT-related solutions) and this classification can be refined by a prioritization of the ideas via the Payoff-Matrix (e.g., Quick Win).

Step 3.2: Fig. 3 exemplifies the subtypes of complementary interdependencies on a meta model level. Considering the *subtype A*, the meta models of techniques (e.g., KANO Model and Driver Tree) share identical classes. This is indicated by the dotted line "#2", hinting at the common class "KANO category" in MM_1 and MM_2.

Concerning the *subtype B* on a meta model level, a relationship exists between core concepts of different techniques (e.g., KANO Model and FMEA), which are represented by dissimilar classes, e.g., "KANO category" and "severity number" (dotted line "#3"). However, the core concepts follow a common purpose (cf. [10]) of transforming input to output information. Regarding *subtype C*, the meta model of a quality technique has one or more classes representing types of output information that are related to classes representing input information types of another meta model. In so doing, the classes representing types of input and output information are not identical as exemplified for the CTQ/CTB-Matrix and the Ishikawa Diagram (see MM_4 and MM_5 – dotted line "#4"). Finally *(subtype D)*, core concepts of quality techniques supporting the transformation of input information to output information, e.g., by prioritizing or categorizing information, may cause a complementary interdependency. Such concepts, e.g., "payoff category", are explicitly represented by separate classes in the meta models and implicitly become evident in the labels of the classes representing types of output information. Further, identical classes representing types of input information are given (e.g., improvement idea). Fig. 3 and the dotted lines "#5" demonstrate this particular form of synergy (MM_6 and MM_7).

Step 3.3: Based on these findings, four indicators of complementary interdependencies on a meta model level, numbered #2 to #5, were derived:

Table 2. Indicators of complementary interdependencies on a meta model level

Indicator #2 *(subtype A):* Two *quality techniques i and j* have a complementary interdependency in case their corresponding meta models (MM_i and MM_j) share one or more identical classes representing common core concepts serving a particular purpose (e.g., prioritization) during the transformation of input to output information. The affected classes represent neither output nor input information types on a meta model level.
Indicator #3 *(subtype B):* Two *quality techniques i and j* have a complementary interdependency if the corresponding meta models (MM_i and MM_j) contain classes representing concepts serving a common purpose during the transformation of input to output information in a particular project (e.g., classification of information) but are not identical per se. In that context, the output information created by applying *technique i* creates knowledge that facilitates the use of *technique j*.
Indicator #4 *(subtype C):* Two *quality techniques i and j* with the corresponding meta models (MM_i and MM_j) have a complementary interdependency, if the output information represented by classes in MM_i on a type level facilitates the specification of input information for *technique j*, represented by classes indicating types of input information in MM_j. However, the output information produced by *quality technique i* is no mandatory prerequisite for applying *quality technique j* in a project.
Indicator #5 *(subtype D):* Two *quality techniques i and j* have a complementary interdependency, if the corresponding meta models (MM_i and MM_j) contain classes representing particular concepts to transform input to output information, whereas the general purpose of the concepts (e.g., prioritization, classification) varies for the *techniques i and j*. The nature of these concepts becomes evident by the classes representing types of output information in the meta models MM_i and MM_j. The combined use of the techniques allows to reflect results from complementary perspectives (e.g., classified and prioritized improvement ideas). MM_i and MM_j share identical classes for representing types of input information.

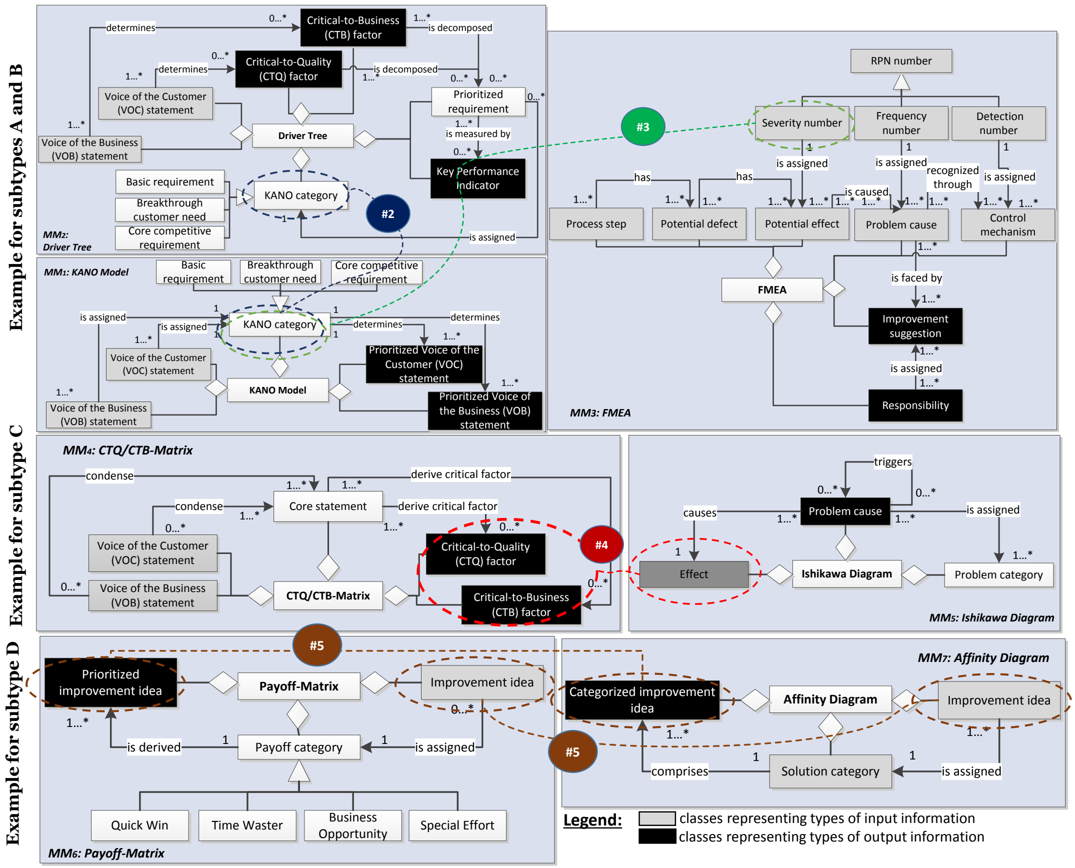


Figure 3. Example for complementary interdependencies (indicators #2 to #5)

5.3 Summary of Substituting and Rivalling Interdependencies

As mentioned, beneficial synergies between quality techniques are primarily determined by conditional and complementary interdependencies. However, further types exist (see Table 1) for which we will briefly summarize the results.

Substituting interdependencies are given in case two or more quality techniques follow the same purpose and produce an identical type of output information (e.g., improvement ideas) (cf. [11]). Substituting interdependencies were given for those quality techniques in our sample that shared a common purpose (e.g., identification of problem causes) and aimed at the production of identical types of output information. An example would be the interdependency between the Measurement-Matrix and the Driver Tree (cf. [12], [17]), with both techniques sharing the purpose of defining KPIs. Accordingly, on a meta model level, substituting interdependencies become obvious by common output information type classes of the techniques. The classes representing types of input information may be different though because the information processed by the techniques may vary. The following indicator was thus derived:

Indicator #6: Two *quality techniques i* and *j* have a substituting interdependency in case the corresponding meta models (MM_i and MM_j) have identical classes representing types of output information, with the *techniques i and j* sharing a common purpose within a project. The classes representing types of input information may be identical or different from one another.

Rivalling interdependencies exist, in case the combined application of quality techniques might lead to results that are contradictory to one another (cf. [11]). On a meta model level, rivalling interdependencies cannot be identified unambiguously. Generally, the types of output information generated are decisive, which is similar to substituting interdependencies. Thus, there is the danger of producing identical types of output information (e.g., improvement ideas), which, however, may contradict each other (cost-oriented vs. customer-oriented improvement ideas) (cf. [11]). Nevertheless, an unambiguous characterization by means of classes of a meta model cannot be done and thus no indicators were derived.

In summary, six indicators were defined pointing to FIs on a meta model level allowing to identify synergies between techniques. No further indicators or subtypes of a FI, allowing to specify the interplay between techniques in a generally valid manner, were found by the researchers performing steps 3.1 to 3.3. More, any interplay between quality techniques in the sample set could be expressed by the indicators as introduced.

6 Proof of Concept (Step 4)

The applicability as well as the usability of the indicators were to be validated. For that purpose, we created a prototype of the aforementioned modeling tool (see section 3) building on FIs in a first step. The prototype served as a proof of concept (cf. [24]) evidencing that the concept of indicators of FIs as well as the corresponding meta models could be realized in the form of an executable software tool. This was important considering the feasibility of the aforementioned DS project (see section 3).

In a second step, a usability study was conducted, reverting to the prototype and two case studies, to prove the beneficial impact of FIs between quality techniques on the development of process improvement suggestions.

Our prototype contains the 14 quality techniques of the sample set that were most frequently applied by the said automotive bank *(tool download as an MS Windows installation package: http://tinyurl.com/zc9rpnp)*. The quality techniques were realized as model types based on the corresponding meta models (see section 4.2). The receipt of an executable prototype demonstrated the validity of these meta models once again in terms of consistency, syntactical correctness and completeness. The model types could be used by project participants straight away for creating results and documenting outcomes in PQM initiatives. The indicators, which specify beneficial interdependencies between the techniques, enabled the development of algorithms that either allow the user to automatically transfer results between particular quality techniques (conditional interdependencies) or to specify and refine the outcomes of a technique by using complementary techniques in addition (complementary interdependencies). In the prototype, conditional interdependencies are automatically exploited whereas complementary interdependencies can be drawn upon optionally. Fig. 4 gives an example for conditional interdependencies. The model on the left shows an excerpt of the CTQ/CTB-Matrix designed as a model type, codifying customer statements (VOC statements) stemming from a project to improve the document management process at the aforementioned automotive bank. From these, the CTQ factor *"reduction of cycle time (…) to two working days"* is derived. Because of a conditional interdependency (indicator #1), the CTQ factor as defined is automatically referenced by the Measurement-Matrix straight away without the data having to be re-entered from the user side. Accordingly, KPIs such as the "overall cycle time" or the "number of errors in archiving" are specified to measure the goal achievement.

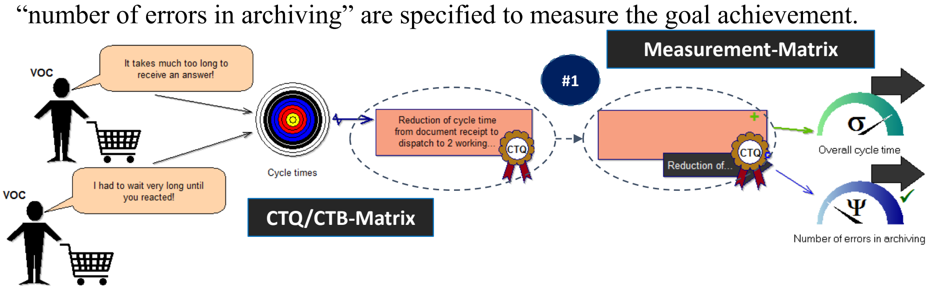


Figure 4. Example from the prototype

An example of a complementary interdependency would be the use of the "payoff categories" in the Affinity Diagram (see the following link for a detailed example: *http://tinyurl.com/zbctxt8*). As an additional proof of the beneficial impact of FIs in terms of project conduction, a usability study of the prototype by means of the SUMI questionnaire (Software Usability Measurement Inventory) *(http://sumi.ucc.ie/)* was performed with 32 Master degree students of a German university. The SUMI questionnaire is a well-established standard for measuring software usability and builds on the five dimensions "efficiency", "affect", "helpfulness", "control", and "learnability" [34]. The intention was to investigate whether users perceived the synergetic use of quality techniques (with complementary and conditional interdependencies) as beneficial for developing process improvement ideas or not. For that purpose, two case studies were drawn upon. The first case study was based on a real life project conducted in cooperation with the mentioned automotive bank to improve the end-of-terms process. The second case study described a fictitious check-in process at an airport. 17 students worked on the first case study and 15 dealt with the second one. Based on given problem statements, the students were supposed to develop suggestions to overcome process weaknesses using the prototype *(material download and detailed results of the study: http://tinyurl.com/zb5r5lk)*. Afterwards, they were asked to fill out the SUMI questionnaire, and an aggregated usability rating was received across both case studies. Taking into account two case studies enabled a more nuanced assessment of usability, because the results were not imprinted by one particular scenario only. In our context, the ratings for the dimension "efficiency" were of particular interest, because it captures the degree as to which the software supports a user in conducting tasks (cf. [34]). Considering the reference score of "50" as proposed by the SUMI reference database (cf. [34]), the users felt well supported by the prototype in deriving improvement suggestions, as indicated by the efficiency ratings *(mean: 51,56; median: 53)*. Making use of the references between the quality techniques, due to beneficial FIs, definitely had a huge share in that, which becomes evident by the SUMI item consensual analysis [34] and the user comments. Users highly appreciated the tool's functionality to easily move from one task to another and, hence, to further refine or process results by using beneficial combinations of techniques (see also: *http://tinyurl.com/zb5r5lk*).

7 Discussion

First, referring to our RQ, it was shown that meta modeling allows researchers to precisely assess the essence of quality techniques and to identify FIs on the base of indicators, respectively. The indicators foster knowledge creation regarding those quality techniques that may be beneficially combined in a project and, thus, help to develop methodological support for quality initiatives. Hence, synergetic quality techniques can be logically arranged and integrated by method or software engineers to come to an enterprise-adapted PQM method (cf. [13], [35]) that meets a firm's specific needs (cf. [9]). In this respect, the combined use of synergetic quality techniques backs the goal-oriented elicitation of process knowledge and its transformation into improvement opportunities, with process knowledge being a key factor to influencing project success (cf. [7]). Considering the toolbox of the automotive bank, which was part of our sample set, we found all FIs as introduced in Table 1 with the corresponding subtypes, which we were able to delineate in this study. However, especially conditional and complementary interdependencies were encountered particularly often.

Second, software support for practitioners for systematically eliciting, documenting and communicating results in the course of a PQM project – even across company borders – can be established as indicated in sections 3 and 6. On that base, project data may be further analyzed by means of reports promoting the querying and capitalization of process-related knowledge generated in PQM projects (cf. [36]). In the paper at hand, the indicators on a meta model level, specifying beneficial FIs, enabled to create algorithms for the automated transfer of project data between quality techniques to be further processed (see section 6).

However, as a restriction, meta modeling requires particular skills and knowledge from the user side. Thus, identifying and exploiting FIs on the base of indicators on a meta model level is an approach, which is most likely interesting for method or software engineers but probably only of little interest to employees who are less of an expert.

8 Conclusion

This research dealt with the question of how to purposefully identify FIs between quality techniques by reverting to their meta models. We learned that meta models, widely established in IS research, are helpful to explicate the functionality of quality techniques and to generally explain the occurrence of FIs. However, there are limitations to this study: the set of techniques analyzed for the derivation of indicators was restricted to the toolbox applied at an automotive bank and to those techniques derived from literature. Thus, further subtypes of FIs may potentially exist, building on fruitful combinations of techniques neither literature nor practice is yet aware of.

Therefore, further research to consider additional techniques is required, although our sample set comprises techniques that are widespread in practice and frequently used in PQM projects across different industries (cf. [10], [12]), which contributes to the general validity of the results. Generally, subjectivity cannot be entirely excluded considering the derivation of FIs as well as the definition of indicators. However, by the discussion between researchers and the consolidation of results, subjective imprints were to be minimized as far as possible. Nonetheless, further indicators for subtypes of FIs not unveiled by this research may occur. Yet, the indicators proved suitable to fully explain the interplay between techniques in the sample set at hand. In future, the indicators will be used for developing guidelines for practitioners explicating which techniques can be beneficially used in combination. More, the prototype will be revised considering practitioners' feedback and designed to run on different platforms.

Appendix

The following list presents the sample set of the quality techniques of our study. Due to page restrictions, the FIs are only exemplified for four techniques. More details on the techniques and interdependencies can be found at: *http://tinyurl.com/zbctxt8*

<table>
<tr><th>Technique</th><th>FI with other techniques (no.)</th><th>Technique</th><th>FI with other techniques (no.)</th></tr>
<tr><td>**1) CTQ/CTB-Matrix**</td><td>*CoI:* supported by no. 3, 5</td><td>18) Brainwriting</td><td>*CoI:* supports no. 16; supported by no. 9, 10, 11, 12, 13, 14, 15;
SI: substitute for no. 17, 19, 21, 22</td></tr>
<tr><td>2) Driver Tree</td><td>*CoI:* supported by no.3;
SI: substitute for no. 7</td><td>19) RAMMPP-Matrix</td><td>*CoI:* supports no. 16; supported by no. 9, 10, 11, 12, 13, 14, 15;
SI: substitute for no. 17, 18, 21, 22</td></tr>
<tr><td colspan="2">3) KANO Model, 4) Process Map, **5) SIPOC Diagram**, 6) Project Charter, **7) Measurement-Matrix, 8) Data Collection Plan, 9) Descriptive Statistics, 10) As-Is Process Modeling, 11) Value-Stream-Map, 12) Ishikawa Diagram**, 13) Relation Diagram, **14) FMEA**, 15) Time Analysis, **16) Should-Be Process Modeling, 17) Brainstorming**</td><td colspan="2">**20) Affinity Diagram**, 21) SCAMPER Technique, 22) Lean for Service, 23) Place Cipher Approach, 24) Prioritization-Matrix, 25) Cost-Benefit Analysis, 26) Payoff-Matrix, 27) Town Meeting Worksheet, 28) Roll Out Plan, **29) Process Documentation, 30) Reaction Plan**, 31) QFD, 32) SERVQUAL, 33) Service Quality Map, 34) Design of Experiments</td></tr>
<tr><td colspan="4">Techniques no. 1 to 30: quality techniques derived from the toolbox at the automotive bank; no. 31 to 34: derived from literature; techniques considered by the prototype are printed in "bold".
Legend: *CoI:* complementary interdependencies; *SI:* substituting interdependencies; *CdI:* conditional interdependencies; *RI:* rivalling interdependencies</td></tr>
</table>

References

1. CapGemini, www.de.capgemini.com/it-trends-studie (Accessed: 17.08.2016)
2. Harmon, P.: The State of Business Process Management – 2016. BPTrends (2016)
3. Stracke, C.: Process-oriented quality management. In: Ehlers U-D, Pawlowski JM (eds.) Handbook on Quality and Standardisation in E-Learning, pp 79-96. Springer (2006)
4. Harmon, P.: Business Process Change. 2nd edition. Elsevier LTD, Oxford (2007)
5. Chakravorty, S.S.: Where Process-Improvement Projects Go Wrong. Wall Street J. (2010)
6. Erdani, Y., Hunger, A., Werner, S., Mertens, S.: Ternary grid as a potentially new technique for knowledge elicitation/acquisition. In: IEEE Conference on Intelligent Systems (2004)
7. Seethamraju, R., Marjanovic, O.: Role of process knowledge in business process improvement methodology: a case study. BPM Journal 15, 920-936 (2009)

8. Gijo, E.V., Rao, T., S.: Six Sigma Implementation – Hurdles and More Hurdles. Total Quality Management & Business Excellence 16, 721-725 (2005)
9. Davis, D.: 3rd Biennial PEX Network Report: State of the Industry – Trends and Success Factors in Business Process Excellence (2013)
10. Hagemeyer, C., Gershenson, J., K., Johnson, D., M.: Classification and application of problem solving quality tools: A manufacturing case study. TQM Mag. 18, 455-483 (2006)
11. Bruhn, M.: Operative Gestaltung des Qualitätsmanagements für Dienstleistungen. In: Bruhn, M. (ed.) Qualitätsmanagement für Dienstleistungen, pp. 251-354. Springer (2013)
12. Meran, R., John, A., Roenpage, O., Staudter, C.: Six Sigma+Lean Toolset. Springer (2013)
13. Fill, H.-G., Karagiannis, D.: On the Conceptualisation of Modelling Methods Using the ADOxx Meta Modelling Platform. EMISA – An International Journal 8, 4-25 (2013)
14. Seidewitz, E.: What models mean. IEEE Software 20, 26-32 (2003)
15. Gutzwiller, T.A.: Das CC RIM-Referenzmodell für den Entwurf von betrieblichen, transaktionsorientierten Informationssystemen. Physica-Verlag, Heidelberg (1994)
16. Ishikawa, K.: Guide to Quality Control. Tokyo (1980)
17. Hutwelker, R.: Hard und Soft Facts zur Prozessverbesserung – Six Sigma. Report (2005)
18. Kleppe, A.G., Warmer, J.B., Bast, W.: MDA explained: the model driven architecture: practice and promise. Addison-Wesley Professional (2003)
19. Anaby-Tavor, A. et al.: Insights into enterprise conceptual modeling. Data & Knowledge Engineering 69, 1302-1318 (2010)
20. Fill, H.-G. et al.: An Approach to Support the Performance Management of Public Health Authorities using an IT based Modeling Method. 10th Wirtschaftsinformatik, Zürich (2011)
21. Amaravadi, C.S., Lee, I.: The dimensions of process knowledge. Knowledge and Process Management 12, 65-76 (2005)
22. Dodgson, M.: Organizational learning: a review of some literatures. Organization Studies 14, 375-394 (1993)
23. Nonaka, I.: The knowledge-creating company. Harvard Business Rev. 85, 162-171 (2007)
24. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28, 75-105 (2004)
25. Rothchild, I.: Induction, deduction, and the scientific method. Soc. study Reprod. (2006)
26. Johannsen, F.: State of the art concerning the integration of methods and techniques in quality management. In: 19th European Conference on Inf. Sys. (ECIS), Helsinki (2011)
27. Fraser, M.D., Kumar, K., Vaishnavi, V.K.: Strategies for incorporating formal specifications in software development. Communications of the ACM 37, 74-86 (1994)
28. Fill, H.-G., Redmond, T., Karagiannis, D.: FDMM: A Formalism for Describing ADOxx Meta Models and Models. In: Proceedings ICEIS, Poland (2012)
29. Batini, C., Lenzerini, M., Navathe, S.B.: A Comparative Analysis of Methodologies for Database Schema Integration. ACM Computing Surveys 18, 323-364 (1986)
30. Kano, N., Seraku, N., Takahashi, F., Tsuji, S.: Attractive Quality and Must-be Quality. Hinshitsu: The Journal of the Japanese Society for Quality Control 14, 39-48 (1984)
31. George, M.L. et al.: Lean Six Sigma Pocket Toolbox. McGraw-Hill, New York (2005)
32. Shahin, A.: Integration of FMEA and the Kano model: An exploratory examination. International Journal of Quality & Reliability Management 21, 731-746 (2004)
33. Ulrich, D., Kerr, S., Ashkenas, R.: The GE Work-Out. McGraw-Hill, New York et al. (2002)
34. Kirakowski, J.: SUMI user handbook. HFRG, University College Cork, Ireland (1998)
35. Ralyté, J., Rolland, C.: An Approach for Method Reengineering. In: International Conference on Conceptual Modeling, pp 471-484. Springer (2001)
36. Xu, Y., Bernard, A., Perry, N., Lian, L.: Managing knowledge management tools: a systematic classification and comparison. In: Int. MASS Conf., pp 1-4. IEEE Press (2011)

From Doves, Magpies and Urban Sparrows: The External Strategic Forces Driving Knowledge Protection Approaches in German SMEs

David Nienaber[1,2,3]

[1] TU München, Chair of Corporate and Intellectual Property Law, Munich, Germany
David.Nienaber@tum.de
[2] Deloitte GmbH, Risk Advisory, Munich, Germany
dnienaber@deloitte.de
[3] BluePaaS Consulting, Munich, Germany
David.Nienaber@BluePaaS.de

Abstract. Knowledge protection serves as the means to securing valuable knowledge assets, the basis of firm strategy according to the resource-based view. Current research has not taken into account external strategic drivers. This study focuses on identifying these strategic drivers and identifies related patterns in protection approaches. I employ an explorative qualitative research design, including 9 German SMEs. The results show that firms can be categorized into 3 protection approaches. Firms set up protection like Doves, Magpies or Urban Sparrows. They are peaceful, protecting their most-valuable assets or invest in general knowledge protection, respectively. The results show that the approach depends on the external strategic drivers of an industry. Consequently, the external perspective should be included in future research. Managers can use the insights to set up knowledge protection in a focused and strategy-aligned manner.

Keywords: knowledge management, knowledge protection, knowledge security, industry-oriented view, resource-based view

1 Introduction

Knowledge is an important source of competitive advantage [1]. Its protection strives to ensures its sustainability [2]. However, what are the external drivers that make protection necessary and is protection of competitive advantage the only reason? Research on protection focuses on knowledge as a source of competitive advantage, the internal strategic perspective known as the resource- and knowledge-based view (RBV and KBV, respectively)[2, 3]. Thus, researchers focused on the mechanisms for protecting knowledge [4] and mechanism effectiveness [5, 6]. Effectiveness is influenced by firm-context and characteristics [7]. Therefore, knowledge protection is more than generalized protection [8].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Nienaber, D. (2017): From Doves, Magpies and Urban Sparrows: The External Strategic Forces Driving Knowledge Protection Approaches in German SMEs, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 731-745

As research focused on the internal perspective, our understanding of external drivers is limited to date. Once a firm has decided to enter competition in an industry, it has to adapt itself to this environment [9]. The goal of this study is to shed light on the external strategic drivers of knowledge protection. Specifically, I aim to understand a) the protection of competitive advantage from the industry perspective and b) the influence of industry characteristics on protection strategy. Porter's five forces, the industry-oriented view (IoV), provide the theoretical lens [9, 10]. I use an explorative qualitative study analyzing the protection approaches of 9 German SMEs [11].

This paper contributes insights how external strategic drivers lead to the protection types, the Doves, Magpies and Urban Sparrows, identified in this paper. First, external strategic drivers relating to competitive advantage, (1) competitive position and (2) knowledge-based entry barriers refine the need for protection derived from the RBV. Furthermore, I discuss that the impact of knowledge loss is more significant relating to competitive positioning as opposed to entry barriers. Regarding knowledge diffusion the impact is significant for competitive positioning and entry barriers. Second, external strategic drivers beyond competitive advantage, (3) protection itself as an entry barrier, (4) entry barriers through regulation and (5) protection-related cost-shifts affect the protection type. Third, the paper shows that a firm's protection type focuses on a firm's owned knowledge but also on knowledge that other firms entrust to it. As these three aspects lead to the variation in protection types, researchers and practitioners should consider all three aspects when defining and evaluating protection approaches.

The remainder of this paper is structured as follows. Section 2 elaborates on the research gap, introduces the industry-oriented view and details knowledge protection as perceived in this paper. Section 3 describes the explorative qualitative research design. Section 4 details our findings on the types of knowledge protection and external strategic drivers. Section 5 discusses the interrelation between external drivers and protection types for owned and entrusted knowledge. Section 6 concludes the paper.

2 Extant Literature and Theoretical Background

This section focuses on the connection between knowledge protection and firm strategy. Section 2.1 shows that the external perspective complements the RBV. As protection should align with strategy and research has focused on the RBV, this paper focuses on the external drivers. Section 2.2 details Porter's framework that focuses on strategy in general. Section 2.3 presents related aspects of knowledge protection.

2.1 External strategic drivers – Complementing the resource-based view on knowledge protection

This section shows how the external strategic perspective helps a firm use its competitive advantage to define competitive position and adapt to its environment. Knowledge plays an important role in this strategic process [1, 3]. Therefore, investigating the external aspects of knowledge protection should complement the current focus on internals.

A protection approach should align with a firm's strategy. The internal and the external strategic perspective both shape that strategy. The external aspects provide the boundaries for a firm. Specifically, external aspects determine how competitive advantage can be used to position a firm relative to its competitors, i.e. a cost-based or a differentiation strategy [9]. As the internals and externals complement each other, both together explain firm performance better than standalone [12]. Internals refer to exploiting strengths and coping with weaknesses whereas externals refer to how a company deals with opportunities and threats [13]. Taking the internal perspective, it is the RBV that focuses on the strengths and weaknesses, specifically on how competitive advantage is derived from a firm's capabilities and strategic resources, especially knowledge [14]. Taking the external perspective, the choice of industry sets not only the threats and opportunities but also the boundaries under which a firm adapts to its environment [9]. Both perspectives are connected. A firm's strategic position in an industry is determined relative to its competitors. The basis is comparing a firm's advantages with competitors [9]. Porter's cost-based and differentiation strategies are such strategic positions. On the one hand, certain advantages lead to lower costs, e.g. due to a superior production technique. On the other hand, advantages lead to differentiating features like superior designs that competitors cannot offer. Many of the advantages are based on exclusive knowledge [1]. The decision which competitive advantages to foster illustrates how a firm can use its knowledge to derive an industry-oriented strategy, a position relative to its competitors. Depending on the chosen strategy, some advantages and related knowledge become important others obsolete.

Extant research on knowledge protection follows the notion of knowledge as a driver of competitive advantage, the internal strategic perspective. The vast majority of studies focuses on ways to protect competitive advantage by protecting the underlying knowledge. Desouza summarized the state of research in the protection of assets as the gap between knowledge management and information security [15]. Addressing the gap in his studies, the link to knowledge as a driver of competitive advantage remains the underlying assumption [16, 17]. In a recent structured literature review, Manhart and Thalmann identify 4 research streams that are all fiercely related to the RBV [4]. Though some studies cope with external aspects of the firm, for example protecting knowledge in partnerships, securing competitive advantage remains in the focus.

In order to complement the internal perspective, this paper explores the externals of protection by asking how external strategic drivers shape protection in SMEs. In detail, the paper focuses on strategic drivers related to knowledge-based competitive advantage and drivers beyond as well as their impact on a firm's protection approach.

2.2 The Industry-oriented View – Providing the Theoretical Lens

This paper uses Porter's five forces framework to analyze the external strategic drivers of a firm [10]. The framework focuses on general strategic aspects of a firm, not on protection itself. As protection should align with strategy, I use the framework to identify the strategic drivers and relate them with knowledge protection in the findings and discussion section. The external drivers are determined by vertical competition, horizontal competition and the threat of substitutes [9, 10]. Along the horizontal dimension, a company may use competitive positioning based on cost or differentiation, set up entry barriers or use signaling to fight industry participants. Along the vertical dimension, suppliers and customers may demand extended services and thereby shift the costs of performing these services. Furthermore, they may integrate forward or backward in the value chain, respectively. Entry barriers can prevent integration. I do not discuss the threat of substitutes as it did not provide any insights from the data.

The horizontal dimension focuses on existing competitors and potential new entrants into the industry. Existing competitors are fought by defining competitive strategies, i.e. cost-based or differentiation-based competition [10]. Both strategies are based on the competitive advantages of a firm. The specific advantages are selected and put together into a single competitive strategy. Furthermore, a firm can fend off new entrants by building entry barriers and by signaling [10]. Entry barriers are advantages of established firms in comparison to new entrants. They can take the form of competitive advantages that the entrants need before entering competition or regulations to comply with. Signaling describes messages that firms send to new entrants about their reactions [10]. A firm may say, if you enter our market we will cut prices dramatically. Thereby, the industry may become unattractive for the entrant.

The vertical dimension focuses on a firm's customers and their suppliers. Vertical firms may shift costs to other players or perform integration [10]. Integration means that firms on another level of the value chain enter competition with the company. Suppliers may integrate forward; customers may integrate backward. Thereby, they extend their value chain by offering similar goods and services. To fend off integration behavior, firms need to set up entry barriers. Furthermore, the shifting of costs means that suppliers or customers demand the company to execute additional activities. Therefore, the costs of performing the activities are shifted to the firm.

2.3 Knowledge Protection, Implementation and SMEs

The paper views protection as strategy and means to protect knowledge from loss [18], unintended diffusion [19] and to reduce its visibility [20]. In contrast to knowledge security, the study excludes knowledge being altered, becoming obsolete or unavailable [8]. Knowledge is an intangible asset. It is embedded in people's minds, tacit knowledge [3]. It can be documented to store and share it, codified knowledge [17]. Codified knowledge shows characteristics similar to information assets. Thus, information security is a part of knowledge protection with a focus on codified knowledge [15]. However, knowledge protection focuses on all knowledge assets.

Protection mechanisms are formal or informal [7, 21, 22]. Formal mechanisms are intellectual property rights, contracts and labor legislation. Informal mechanisms take the form of secrecy, lead-time and tacitness. The mechanisms are often used in combination [23]. In order to implement mechanisms, specific measures are used [4]. They are grouped into legal, organizational, and technical measures [5, 6, 24]. Legal measures refer to contracts like non-disclosure-agreements [6] and intellectual property rights [5], for example. Organizational measures refer to recruiting [17], training [17], organizational culture [3] and alike. Technical measures refer to means by which access to knowledge is protected, e.g. by firewalls, passwords or physical security [4].

The focus on SMEs prays for attention to informal mechanisms. SMEs are resource-constrained. Thus, they favor informal mechanisms as they come at lower costs [21]. Especially secrecy plays an important role compared to formal mechanisms such as patents [22]. Additionally, protection in SMEs focuses on owned and entrusted knowledge. The latter is important as SMEs operate relatively large firm networks. This helps SMEs allocating resources to core activities [21]. Such a network requires vast knowledge exchange. Thus, firm's entrust their knowledge to the SMEs and vice versa.

3 Method

I employ a qualitative, explorative research design. The design is applicable, as external factors are not yet well understood [11]. The study focuses on German SMEs. Located in a knowledge society, they understand the importance to protect knowledge. Extensive variation in external factors is expected as in most other aspects of SMEs. SMEs are defined following the IfM Bonn [25], excluding blue collars from head count.

Firms were identified using theoretical sampling based on firm characteristics [26]. Seek maximum variation [27], I added firms until the results appeared mature and validated. Firms had to show high variance in characteristics that likely impact knowledge protection, size, industry, service orientation, R&D intensity, competition, network complexity and legal risk. Thereby, I ensured retrieval of rich data. Participants were chosen as they a) had personal interest, b) perceived threats to their firm or c) expect their firm to be confronted with this topic. In summary, I expect deep insights and high variation in external strategic drivers and protection types.

The study investigates the firm level using interviews and additional data as the main sources. 34 interviews were conducted with 14 managers of 9 firms. Three steps were taken to gather firm-level data via the interviews. First, all questions specifically address the firm, not the individual's opinion. Second, multiple interviewees were conducted in 5 firms. Triangulation between interviewees ensures further traceability to the firm level. Third, participants had insights in the firm's strategic dealings and protection approach due to their positions, managing director, CIO, Head of Business Development, Heads of Sales and Head of Marketing. On average 3.8 interviews per firm were conducted. The duration of interviews ranges from 0.5 up to 3 hours.

Knowledge protection approaches are highly sensitive to the firms. Recording was not allowed in 79% of the interviews. Thus, I defined a strategy to ensure data validity using semi-structured guides, active listening and triangulation. The semi-structured guides allowed for extensive note taking. Active listening served as the immediate validation step to retain the original voice by mirroring the information to the interviewee. All notes were reviewed with the participant after each section. For final validation, follow-up calls and meetings were used in situations where doubt remained.

Triangulation of interview data with firm-level data sources ensured data validity. The sources comprise 67 public and private documents. Public documents are brochures, websites, studies, financial reports, product brochures, corporate videos, general terms and conditions as well as newspaper articles. Private documents are auditor reports, non-disclosure-agreements, code-of-conducts, internal policies and brochures on compliant behavior.

Coding and analysis focused on the external factors, identified along the vertical and horizontal dimensions [10]. I extracted the codes to a new document for focused analysis (see [28] for a details). Along each dimension, distinct themes emerged (c.f. section 4). The relative number of code segments per theme is stored in a firm-vector (x_i,y_i,z_i) with dimension i. A theme was assigned, if it scored 30% or more. Beyond pure types, mixed types with two or more themes along a dimension and themes that vary across dimensions emerged. Expert interviews served as final validation.

4 Findings

This chapter presents the external drivers and protection types. The results reveal systematic variation along the horizontal and vertical dimension. The variation leads to three distinct themes for each dimension, based on the external drivers (c.f. Table 1). Drivers and themes result in the Dove, Magpie and Urban Sparrow protection type (c.f. Table 1). I begin with a summary of the themes, followed by a detailed illustration of the protection types. The external strategic drivers are discussed in section 5.

Along the horizontal dimension, firms take different approaches to fend off competitors and new entrants. Specifically, firms a) underestimate threats and perceive protection as an insignificant matter while relying on informal mechanisms on an as it happens basis; others are aware and react by b) focusing protection on their most valuable assets, their crown-jewels, or by c) implementing a broad protection that covers all knowledge, whether valuable or not. The important difference is that broad protection tries to prevent a breakdown of operations, which is a knowledge security topic, whereas the crown-jewel approach tries to stop knowledge loss and theft.

Along the vertical dimension, firms cope with threats from suppliers and customers, pursuing different approaches as well. Either, they a) misinterpret the threats, or b) protect entrusted knowledge, which is knowledge that a third party provides to the firm, and owned knowledge based on project protection requirements, or c) follow a broad protection approach. In essence, firms extend protection approaches aimed at their own assets to the assets of third parties, especially customers, and vice versa.

4.1 Doves

Doves resemble the first type of protection approaches. Doves perceive the business world as peaceful. They believe in fair business practices and protection is a matter of little significance to them. Doves as the symbol of peace do not get intentionally into struggles with competitors or other third parties. When forced into a struggle, Doves solely care about their own assets. Here they rely on informal protection mechanisms or formal protection mechanisms that come at little cost, such as contracts. Small firms often behave like Doves. They do not perceive competition as intense. Their network is small. Usually, they do not do international business and do not face legal risks (c.f. Table 2). The themes along the dimensions illustrate the Dove-view on protection.

Table 1. The variance along strategic dimensions that drive knowledge protection

Strategic Dimensions	*Variance in Meanings* *Doves*	*Magpies*	*Urban Sparrows*
Horizontal	*Underestimation & insignificance:* - Underestimation of threats. - Focus on informal mechanisms. - Insignificance as protection seems unnecessary or ineffective.	*Focused protection:* - Protection of crown jewels only. - Knowledge that is at the basis of firm success belongs to the crown jewels.	*Derivation from information security:* - Heterogeneity of knowledge leads to a general broad approach. - General protection of decision systems and production processes. - Regulatory requirements.
Vertical	*Misinterpretation of threats:.* - Informal approach: Personal and reputation-based customer relations. - Perceived threat: Stealing of customers instead of knowledge theft by third parties.	*Project protection requirements:* - Project-based protection of customer knowledge. - Entry barriers by customer-specific protection and trust. - Protection against forward integration of suppliers by tailored approach.	*General protection requirements:* - Broad protection approach for customer knowledge. - Broad protection of supplier and partner portals. - Protection as a part of the firm's product or services.
External Strategic Drivers	***Competitive advantage:*** - Threats from competitors. - Focus on own knowledge. - Ignoring vertical and entry threats.	***Competitive advantage:*** - Vertical and horizontal threats. - Protection of crown jewels. ***Entry barriers:*** - Customer trust. ***Cost-shift – 3rd party requirements:*** - Customer-specific protection.	***Competitive advantage:*** - Vertical and horizontal threats. - Heterogeneity of knowledge. - Focus on securing operations. ***Entry barriers:*** - Customer trust. - Compliance with regulation. ***Cost-shift – 3rd party requirements:*** - General third-party requirements.

Along the horizontal dimension, Doves underestimate the threat from competitors and new entrants. They use informal protection mechanisms. They perceive protection as insignificant as they feel that protection mechanisms are ineffective. Regarding competitor and entrant threats, the case of the event equipment rental company exemplifies how Doves perceive their threat situation. In the firm's business knowledge like product rental margins and optimal product stocking is key knowledge. Discussing its protection with his managing director, the company's IT manager received the following reaction.

What I just wanted to say, when you are discussing this with our managing director… he only says… hm, who should ever attack us?

Doves perceive formal protection measures as unnecessary, relying on informal measures. The electrical engineering company is based on a patent for inductive transmission. After expiration, the technology proved inimitably due to its tacitness, rather by accident. Only the managing director knows it well enough so he can sell it. The engineer's job is to set up the products to a solution based on the technology and given specifications. The business development manager of the company elaborates first on the tacit protection before addressing the critical role of the managing director.

The whole customer-specific area…[…] you just cannot copy it. […] The transfer method, the technology that forms the origin of the company, was a patent. So, it is public now as it is older than 20 years. But … even this is too complicated for people so that they wouldn't even want to imitate it. It is not a technology, which you simply cut open. You can't just look inside and see how it works. […]

So our managing director is responsible for sales and that is just right. So, we had one that wanted to do sales and he left. You have to say, that was unproblematic. He went to another industry…because, how shall I say, the composition, the customer-specific thing is tough. So … he (the customer) asks something and in that moment you have to run through your 100.000 options, because it is not a product, you have to give a solution. […] And of that, no one else is capable of.

Doves even perceive formal and informal protection mechanisms as ineffective. Therefore, the protection is of little significance. The building company provides an example. It uses secrecy in combination with motivational HR measures. It aims to prevent employees leaving. Labor legislation complements the informal mechanisms. Yet, the managing director feels that he is helpless in case key employees leave and take the most important asset, knowledge about customers with them. After elaborating on the HR perspective, he jumps quite agitated to the problem of managers leaving the firm. He even shows some degree of personal insult taken from such events.

*We try to keep our employees happy. We provide them with a car, flexibility on maternal leaves and other things to make them stay. But, this is no assurance that they stay and won't take things with them. […] In that context, you can see a dictum that I much like to use: "The worst scam happens within your innermost circle!" So secrecy is highly important in management, but how can I ever enforce this?**[1]

[1] An asterisk indicates a quote, originating from an interview where recording was not allowed.

Along the vertical dimension, managers do not realize the threat from forward integrating suppliers and backward integrating customers. Their only concern is competitors trying to woo their customers. To prevent customers from leaving, Doves use open communication and build personal as well as reputation-based customer relations. To that end, Doves use informal measures such as employee retention. Thereby, they try to keep employees that have important customer relations. The business development manager illustrates the informal nature of relationship building.

But, let me say it this way. Let's go back one step. Type MAN company xy, ok? The company was served quite well from our side. And now, firm xyz (a competitor) calls, ok? The chance that the MAN manager says ... yes the other one (the competitor) fits well, everything is fine…is really small. Yes, I mean in that regard our customer list is not uncritical when stolen, but…you see.

4.2 Magpies

Magpies are the second type of firms. Magpies like all that glitters, according to folk tale. For a firm, these glittering things are their most valuable assets, crown jewels. They comprise of knowledge owned by Magpies themselves as well as knowledge, which customers entrust to the Magpies. Magpies are rather independent of size. They likely serve a specialized market by a project-driven business model. For example, Magpies implement protection specific to a project in a high legal-risk country (c.f. Table 2). The themes along the dimensions illustrate the Magpie-view on protection.

Along the horizontal dimension, magpies perceive the threat of competitors and new entrants. Magpies derive their competitive strategy from crown jewels. Thus, Magpies focus on protecting these assets. An exemplification of the importance of crown jewels is the following statement by the managing director of an equipment manufacturer. She clearly states what the firm's crown jewels are and how they use tacitness to protect its process technology.

So, let's put it that way. You always have to be ahead. The big know-how is our process technology. Because you cannot see it. But, when you are having a look at the machines, you quickly see how it works. Another topic is automation, all control technology, the controlling, and how a plant is planned and conceptualized. This is another topic, where you can be well ahead of your competitors.

The statement by the CIO of a manufacturer of embedded systems illustrates how the firm derives its competitive strategy from its crown jewel, service delivery. Though he refrains from stating how the related product strategy is protected, this part shows the firm's focus on this important aspect of firm strategy.

*…our focus on Added Services and the planning, which services shall be delivered in the future. This is where we are a big step ahead of our competitors. If they would know about our product strategy, they would enter the market earlier and would start competition much earlier.**

Table 2. Sample characteristics

Characteristics	*Dove*	*Magpie*	*Urban Sparrow*	*Mixed Type*
Number of firms	3	2	2	2
No/partial/Management ownership	0/1/2	0/1/1	0/0/2	1/0/1
Listed stock	0	0	0	1
Low/medium/high size	1/2/0	1/0/1	0/1/1	0/0/2
Service/product/mixed orientation	2/1/0	1/1/0	0/2/0	0/1/1
No/occasional/continuous R&D	2/0/1	1/0/1	0/0/2	0/0/2
Low/medium/high competition	1/1/1	1/0/1	0/1/1	0/0/2
Low/medium/high network complexity	1/2/0	0/2/0	1/0/1	0/0/2
Global legal risks	0	1	2	1

Along the vertical dimension, magpies implement a project-based approach. They implement protection based on customer requirements, building trust. Furthermore, they implement protection measures on a case-by-case basis, preventing forward and backward integration. The managing director of a consulting firm illustrates customer-specific requirements and building trust as follows. He states how his customer's demand protection of specific knowledge that they will share with the firm throughout the engagement.

> *The project documentation contains our methods to a certain extent. And it reflects the customer situation. Moreover, it often contains more sensitive information of the customer, well … also information how a customer approaches specific situations. As this is highly important to his competitors, it has to be protected. Customers use NDAs to protect the project documentation and to put protection requirements on all our activities. […] Customers give their sensitive information to us, so that we can do our work. So we need to build an environment of trust. One main aspect to build trust is an adequate protection of knowledge. […] Customers, especially in the defense & military sector, have levels of security. Here, customers make prescriptions how to handle information of a certain security level.**

Regarding the protection of its own knowledge, Magpies protect their crown jewels in the vertical dimension. Magpies make case-specific decisions about who is provided with which crown jewels. The managing director of the equipment manufacturer describes how the firm gives its crown jewels (see first quote in this section) to its suppliers, according to situation. The decision is project-specific.

> *And, well, … yes, you know it yourself. It is not easy to protect know-how. As said. Reengineering machines, that's what we do every day. You just can't do anything about it. You can only make sure that you are working with suppliers. With which you get along well and… yes, you give them only the information that is absolutely necessary (for project delivery).*

4.3 Urban Sparrows

Urban sparrows invest vast resources for protecting assets, just like the bird. Resources are readily available as the bird lives in a city. Therefore, he can focus on his territorial fights. The territorial fights stand for a broad protection approach against any threat. The firm has resources available and invests in this fight. Large firms with continuous R&D and a complex firm network are likely to employ this approach. Furthermore, global activities lead to high legal risks. Thus, Urban Sparrows cannot rely on legislation for protecting their assts. The themes along the strategic dimensions illustrate the Urban Sparrow-view on knowledge protection.

Along the horizontal dimension, Urban Sparrows implement a broad protection approach that is derived from and a benefit of information security. The reason is heterogeneity of knowledge. The approach focuses on the management decision systems and continuity of production processes, preventing firm breakdown. So, Urban Sparrows tend to derive knowledge protection largely from information security. In addition, Urban Sparrows focus on complying with regulatory standards. The managing director of a wholesaler provides an illustration of heterogeneity. He shows how he uses information security in order to maintain secrecy of codified knowledge. His quote illustrates how heterogeneity of knowledge, as the firm varies in distribution of knowledge assets, leads to the focus on security, rather than focused protection.

> *Relevance of secrecy is quite limited and present in certain parts of the firm, only. It varies from division to division. […] We have a single IT company located in our Group, it is ISO 27001 certified and copes with all IT aspects.*

Protecting decision systems and production processes results in heterogeneity as well. All IT systems, all employees and all customer or supplier knowledge that belong to both or either one system, need protection. The broad protection of knowledge can be described as a benefit of information security measures as well. The following extract of an internal policy shows that secrecy is but a part of the overall approach. Secrecy, as the part relevant to knowledge protection, is only driven by regulation. Again, Sparrows derive protection of codified knowledge from information security.

> *Our data and our IT systems in all technical and business divisions are protected in such a way that expected shutdown periods could be tolerated. […] The requirements of secrecy have a normal level that is oriented towards regulatory compliance. For data of the human resource department, maximum secrecy requirements apply. […]*
>
> *Delayed or erroneous management decision can have widespread consequences. Therefore, the access to current controlling data is highly important when management faces important decisions. For this type of information, a high level of security in terms of availability and integrity has to be guaranteed.*

Along the vertical dimension, Urban Sparrows follow a broad approach as well. Heterogeneity of customer, supplier and partner knowledge, e.g. in IT portals, drive the approach. Even so, knowledge protection can form part of the firm's offerings. A third party manufacturer exemplifies the heterogeneity of customer knowledge. He builds seals based on customer specifications and knowledge.

*As a third party manufacturer, the firm does not just have customer contact data but receives sensitive information like product specifications as well. The specifications are guiding our work. Therefore, it is self-evident that the specifications contain customer secrets.**

The official internal policy illustrates the importance of customer requirements.

As the customer is located in the center of all our activities, the protection of (customer) information from unauthorized access and manipulation is of existential importance for the firm.

Knowledge protection as part of a firm's offerings is driven by customer demand as well. Here, customer demand leads again to a focus on information security. Naturally, these are used to protect codified knowledge. A managing director of a cloud provider, that is active in logistics and production as well, illustrates the logic of applying protection capabilities into other parts of one's firm.

The second topic is our data center, hmm, where in the end only authorized personnel can get in through iris-scanning-systems. These are all topics (knowledge protection), well that have been done back then, as they were customer relevant. Well, you have a point there, so that you can say to the customer, listen up, your data is safe here. Well, in the end it is part of a product. So, it is driven from outside then.

5 Discussion

The protection types show, that protection strategy depends on how a firm wants to and has to protect knowledge. It is important that firms focus on knowledge owned by and entrusted to it. The external strategic drivers affect its protection and set the boundaries. Within the boundaries, firms can decide whether the benefits of protection exceed costs and whether to focus on crown jewels or on all knowledge. Based on the external factors and these two questions, firms adopt a Dove, Magpie or Urban Sparrow protection. In addition, the protection approaches for owned and entrusted knowledge complement or extend each other. The logistics software company, for example uses an Urban Sparrow approach for their customer's knowledge. Yet, they implement additional protection specific for their crown jewels. Figure 1 illustrates and the following paragraphs discuss the external drivers related to and beyond competitive advantage as well as the protection of owned and entrusted knowledge in more detail.

The results reveal external drivers relating to competitive advantage and the RBV. Such drivers determine the protection type along the horizontal and vertical dimension. Here, knowledge needs protection as it either defines a firm's competitive position or builds entry barriers. This differentiation reveals two important insights. First, some firms do not acknowledge the threat posed from all industry participants. Doves only perceive a threat to competitive position and to entry barriers against entrants. However, they neglect the threats posed by integrating suppliers and customers. Thus, the IoV helps firms to fully assess all threats. Second, knowledge loss has a significant impact concerning competitive positioning but has a marginal impact considering entry barriers. From a diffusion perspective both have significant impact. The impact of knowledge loss related to competitive position equals a negative shift relative to competitors. With the competitive advantage gone, competitors are likely to do more business. However, loss at one firm does not alter the barriers of the whole industry. Others will hold up the barrier. Thus, entry gets more likely but the impact on industry attractiveness is marginal. Shifting perspective to diffusion, the impact is significant on position and barriers. Competitors acquiring knowledge that builds a competitive advantage shifts competitive position. The competitor improves his own position, resulting in a negative shift of one's own position. This leads to intensified competition. Similarly, potential entrants use acquired knowledge to overcome entry barriers. For example, a production technique is important to produce at a certain price. Without the technique, products would be too costly and entry into the industry unattractive.

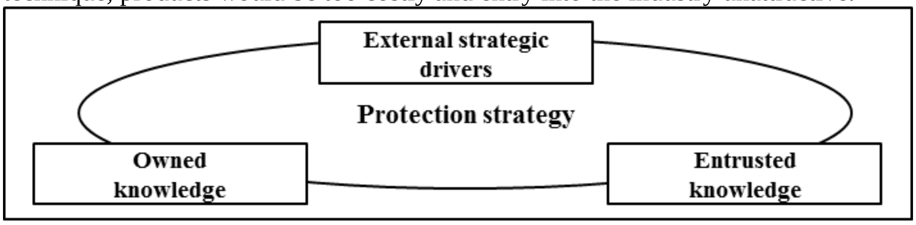


Figure 1: Industry factors and knowledge types as determinants of knowledge protection

The findings present external drivers beyond competitive advantage and the RBV, namely entry barriers relating to protection itself and regulation as well as shifts of protection costs. Such drivers determine the protection type along the vertical dimension. First, protection can build trust of customers and business partners, which serves as an entry barrier. Entrants, forward-integrating suppliers and backward integrating customers, have to implement protection. Otherwise, they cannot enter the market. As such, protection serves as a prerequisite for a business relation. Second, firms have to comply with regulations. While it is natural that firms have to comply with regulations, firms can indeed shape regulations and determine industry standards on the degree of compliance. Depending on this decision, firms can use knowledge protection to set up barriers grounded in regulation. Third, suppliers, customers and business partners may shift costs for knowledge protection on a firm due to their bargaining power. They make another firm protect the knowledge entrust to it. The downside are higher costs, when third parties force a certain degree of protection. The need to protect third party knowledge is a shift of costs on this firm. The upside is a new entry barrier and the option to use the protection level for one's own knowledge.

Protection does not just focus on a firm's owned knowledge but on knowledge entrusted to the firm as well. The Mixed Types show that a firm can protect owned and entrusted knowledge differently. Again, protection depends on the external drivers. Owned knowledge builds strategic position and entry barriers. Entrusted knowledge needs protection to build entry barriers and derive benefits from cost shifts. Thus, protection aims at both, owned and entrusted knowledge. Both aspects interact with each other. The protection mechanisms that are set up for protecting entrusted knowledge can also protect one's own knowledge. It is management's decision in what way to comply with the requirements of industry participants and whether and how to apply these measures to a company's own knowledge.

6 Conclusion

This study has investigated the external strategic drivers of knowledge protection. The drivers were analyzed along the dimensions of Porter's industry-oriented view [9, 10]. The results were structured into themes of protection in the vertical and horizontal industry dimensions. Thereby, the paper identifies external strategic drivers. Thereby, I identified three types of protection, the Doves, Magpies and Urban Sparrows.

In addition, the paper contributes insights into the external strategic drivers of protection. They relate to competitive advantage on the one hand. These drivers are competitive position and knowledge-based entry barriers. On the other hand, they go beyond competitive advantage. Such drivers are protection itself as an entry barrier, entry barriers through regulation and protection-related cost-shifts. Furthermore, the paper calls for considering not only the protection of a firm's own knowledge but also of knowledge that other firms entrust to it. In summary, research and practice should consider these externals when setting up and evaluating protection.

The study has four main limitations, the focus on Germany, the focus on SMEs, the limited amount of interview transcripts and the use of interview data for analyzing organizations. The focus on Germany and SMEs limits generalizability. Firms in other countries act in different environments. Thus, they may encounter further strategic drivers that affect knowledge protection. The focus on SMEs may restrict insights as large firms have more resources available. Thus, protection approaches may not vary as intensively as for SMEs. The third topic limits the validity of our results. In 79% of the interviews, recording was forbidden. The last topic addresses the issue that interview data was used though the object of interest is the firm. The interviewee may be biased towards the firm. Thereby, he may try to shed a firm in a more positive light. Furthermore, interviewees may not report on the situation in a firm but on their general experience. However, I believe that the results are robust due to the measures taken. Future studies should overcome our limitations. Researchers should include firms of all sizes and outside of Germany. In addition, research should try to develop an objective checklist for protection types and conduct interviews with multiple persons per firm.

References

1. Nickerson, J.A., Zenger, T.R.: A Knowledge-Based Theory of the Firm—The Problem-Solving Perspective. Organization Science 15, 617–632 (2004)
2. Wernerfelt, B.: A Resource-Based View of the Firm. Strategic Management Journal 5, 171–180 (1984)
3. Nonaka, I., Toyama, R., Nagata, A.: A Firm as a Knowledge-creating Entity: A New Perspective on the Theory of the Firm. Industrial and Corporate Change 9, 1–20 (2000)
4. Manhart, M., Thalmann, S.: Protecting organizational knowledge: a structured literature review. Journal of Knowledge Management 19, 190–211 (2015)
5. Hertzfeld, H.R., Link, A.N., Vonortas, N.S.: Intellectual property protection mechanisms in research partnerships. Research Policy 35, 825–838 (2006)
6. Norman, P.M.: Are your secrets safe? Knowledge protection in strategic alliances. Business Horizons 44, 51–60 (2001)

7. Olander, H., Hurmelinna-Laukkanen, P., Heilmann, P.: Do SMEs benefit From HRM-Related Knowledge Protection In Innovation Management? International Journal of Innovation Management 15, 593–616 (2011)
8. Bloodgood, J.M., Salisbury, W.D.: Understanding the influence of organizational change strategies on information technology and knowledge management strategies. Decision Support Systems 31, 55–69 (2001)
9. Porter, M.E.: Competitive advantage. Creating and sustaining superior performance. Free Press; Collier Macmillan, New York, London (1985)
10. Porter, M.E.: The Five Competitive Forces that Shape Strategy. Harvard Business Review 86, 79–93 (2008)
11. Glaser, B.G., Strauss, A.L.: The Discovery of Grounded Theory. Aldine Publishing Co, Chicago (1967)
12. Henderson, R., Mitchell, W.: The Interactions of Organizational and Competitive Influences on Strategy and Performance. Strategic Management Journal 18, S.5-14 (1997)
13. Andrews, K.R., Christensen, C.R., Guth, W.D., Learned, E.P.: Business policy: text and cases. Richard D. Irwin, Inc/Irwin-Dorsey Limited, Homewood/Georgetown (1996)
14. Grant, R.M.: The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation. California Management Review 33, 114–135 (1991)
15. Desouza, K.C.: Knowledge Security: An Interesting Research Space. Journal of Information Science and Technology 3, 1–6 (2006)
16. Trkman, P., Desouza, K.C.: Knowledge Risks in Organizational Networks: An Exploratory Framework. Journal of Strategic Information Systems 21, 1–33 (2012)
17. Desouza, K.C., Vanapalli, G.K.: Securing knowledge in organizations: lessons from the defense and intelligence sectors. International Journal of Information Management 25, 85–98 (2005)
18. Jennex, M.E., Durcikova, A.: Assessing Knowledge Loss Risk. In: 46th Hawaii International Conference, S.3478–3487 (2013)
19. Ahmad, A., Bosua, R., Scheepers, R.: Protecting organizational competitive advantage: A knowledge leakage perspective. Computers & Security 42, 27–39 (2014)
20. Lee, S.C., Chang, S.N., Liu, C.Y., Yang, J.: The effect of knowledge protection, knowledge ambiguity, and relational capital on alliance performance. Knowledge and Process Management 14, S.58-69 (2007)
21. Olander, H., Hurmelinna-Laukkanen, P., Mähönen, J.: What's small size got to do with it? Protection of intellectual assets in SMEs. International Journal of Innovation Management 13, 349–370 (2009)
22. Cohen, W.M., Nelson, R.R., Walsh, J.P.: Protecting their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent(or not). NBER Working Paper Series No. 7552, 1–31 (2000)
23. Amara, N., Landry, R., Traoré, N.: Managing the protection of innovations in knowledge-intensive business services. Research Policy 37, 1530–1547 (2008)
24. Bertino, E., Khan, L.R., Sandhu, R., Thuraisingham, B.: Secure knowledge management: confidentiality, trust, and privacy. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 36, 429–438 (2006)
25. Günterberg, B., Wolter, H.-J.: Unternehmensgrößenstatistik 2001/2002 – Daten und Fakten. IfM-Materialen Nr. 157 (2002)
26. Miles, M.B., Huberman, A.M.: Qualitative Data Analysis. Sage, Thousand Oaks (1994)
27. Patton, M.: Qualitative evaluation and research methods. Sage, Beverly Hills (1990)
28. Fauchart, E., Gruber, M.: Darwinians, Communitarians, and Missionaries: The Role of Founder Identity in Entrepreneurship. Academy of Management Journal 54, 935–957 (2011)

Mobiles Lernen für Industrie 4.0: Probleme, Ziele, Lernarrangements

Andreas Engelmann[1], Peter Heinrich[1], Gerhard Schwabe[1]

[1]Universität Zürich, Institut für Informatik, Binzmühlestrasse 14, 8050 Zürich
{engelmann,peterhe,schwabe}@ifi.uzh.ch

Abstract. Bestrebungen im Bereich Industrie 4.0 verändern die Rolle des Menschen in der Produktion nachhaltig. Dadurch verändern sich auch die Anforderungen an die Kompetenzen bei den Mitarbeitern. Anhand des konkreten Fallbeispiels eines Unternehmens im Wandel (Segment Automobilzulieferer), zeigen wir Probleme der Personalentwicklung auf und erstellen konkrete Zielvorgaben an das Kompetenzmanagement. Eine besondere Herausforderung stellt die Transformation von fachspezifischen hin zu fachübergreifenden Aufgabenfeldern dar, ohne dabei langfristig Expertenwissen zu verlieren. Als möglichen Lösungsweg schlagen wir mobile Lernarrangements vor, welche dieses Wissen kapseln und individuell, dezentral und bedarfsgerecht zur Weiterbildung anbieten. Neu erarbeitetes, situatives Wissen erweitert diese Lernumgebung dabei sukzessiv. Der Artikel konkretisiert die Anforderungen an den Kompetenzaufbau im „Industrie 4.0"-Umfeld und leistet auch einen Beitrag zur praktischen Umsetzung.

Keywords: mobiles Lernen, Lernarrangements, Personalentwicklung, Industrie 4.0

1 Einleitung

In den letzten Jahrzehnten sind die Anforderungen an die Qualifikation der Mitarbeiter in nahezu jedem kleinen und mittelständischen Unternehmen, sowie auch in Großunternehmen der Fertigungsindustrie kontinuierlich gestiegen. Man geht davon aus, dass sich dieser Trend durch Entwicklungen in Richtung „Industrie 4.0" noch verstärken wird. Die Produktionsstätten werden dabei grundlegend reorganisiert wobei den Menschen neue Rollen zukommen sollen [1]. Arbeitsabläufe werden flexibilisiert [2] wobei der zukünftige Produktionsmitarbeiter sich insbesondere durch seine Problemlösungsfähigkeiten auszeichnet und im Zusammenspiel mit der Automatisierung eine möglichst hohe Effizienz anstrebt. Ähnliche Konzepte zu Industrie 4.0 werden auch unter den Begriffen „Smart Manufacturing" [3] oder „Advanced Manufacturing" diskutiert [4]. Mitarbeiter werden sich in diesen Umgebungen situationsabhängig neues Spezialwissen in sich fortlaufend verändernden Arbeitsumfeldern aneignen müssen. Das einmalige Lernen von Faktenwissen muss daher einem individuellen „just-in-time" Lernen von Spezialwissen weichen. Dieses Wissen ist aber meist „tacit" oder „internalisiert" (z.B. eigene Erkenntnisse oder

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Engelmann, A.; Heinrich, P.; Schwabe, G. (2017): Mobiles Lernen für Industrie 4.0: Probleme, Ziele, Lernarrangements, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 746-760

Intuitionen) wodurch es schwierig zu formalisieren und auf andere übertragbar ist [5]. Polanyi [6] führte den Begriff des taziten Wissen gar als umfassende Form der Erkenntnis ein, von dem sich nur ein kleiner Teil formalisieren und einfach übertragen lässt. Diese Übertragung kann mit den heute etablierten, oftmals stark standardisierten Lern- und Lehrsettings in Unternehmen nicht befriedigend erreicht werden. Nonaka und Takeuchi [5] bieten das SECI-Modell (Socialization, Externalization, Combination, Internalization) an, welches die kontinuierliche Transformation zwischen implizitem und explizitem Wissen beschreibt [5]. Daraus folgernd müssen Wege gefunden werden, neben dem expliziten Wissen auch das internalisierte (Erfahrungs-) Wissen verfügbar zu machen (Externalization) und individuell oder durch soziale Interaktion weiterzugeben (Socialization).

Mobiles Lernen ist in der Literatur als geeignetes Konzept beschrieben um Wissen individuell und kontextbezogen zu vermitteln [7, 8]. Dabei beschränken sich die Kontexte aber meist auf Situationen der reinen Wissensvermittlung, wie beispielsweise Universitäten [9] oder Museen [10]. Der von der Industrie geforderte arbeitsbegleitende Kompetenzaufbau wurde bisher nur unzureichend beleuchtet. Gerade die technische Fortentwicklung durch Sensorik und Vernetzung erweitern die Umsetzungs-möglichkeiten und bilden die Basis für völlig neue Lernarrangements im „Industrie 4.0“-Umfeld. Die Forschungsfrage lautet daher: „Wo liegen konkrete Potentiale für mobiles Lernen im „Industrie 4.0“-Kontext?“

Anhand des Fallbeispiels eines großen deutschen Automobilzulieferers, der sich gerade in einer Reorganisationsphase der Produktion befindet, konkretisieren wir diese Fragestellung und demonstrieren deren praktische Relevanz. Im Rahmen eines Design-Research-Projektes wollen wir in Zusammenarbeit mit diesem Unternehmen mobile Lernumgebungen „Industrie 4.0“-gerecht gestalten und real einführen. Dieser Artikel fokussiert dabei auf die ersten Phasen des Entwicklungszyklus; der Problemerhebung und Zieldefinition.

2 Grundlagen

Der richtige Umgang mit Wissen und Information ist grundlegend für den Unternehmenserfolg – in jeder Sparte und in jeder Unternehmensgröße. Grant [11] hat Wissen und den Umgang als das grundlegende Konzept für sein Modell „Towards a knowledge based theory of the firm“ benutzt und argumentiert, dass der primäre Zweck einer jeden Firma die Anwendung und Nutzung von Wissen ist. Dabei ist insbesondere zu beachten, dass es um das verkörperte Wissen geht – in den Köpfen der Mitarbeiter gleichermaßen wie auch verkörpert durch Produktionsmaschinen [11]. Grant argumentiert darüber hinaus, dass aufgrund der hohen Komplexität und der begrenzten Aufnahmefähigkeit der Menschen eine Spezialisierung der Arbeitskräfte unumgänglich ist, um das benötigte Wissen an den richtigen Stellen vorzuhalten [11].

Eine breit akzeptierte Definition für „Industrie 4.0“ lässt sich zum gegenwärtigen Zeitpunkt noch nicht finden. Die bestehenden Definitionsversuche sind bislang sehr komplex und beziehen sich meistens auf das Thema der Produktionsautomatisation, obwohl der Themenkomplex auch organisatorische und soziale Bereiche einschließt [12]. Für die Zwecke dieses Artikels legen wir Folgendes fest:

Als Industrie 4.0 verstehen wir technologieunterstützte Reorganisationen industrieller Arbeitsumgebungen mit dem Ziel einer menschzentrierten Selbstorganisation der Abläufe. Diese Sichtweise ist kompatibel mit den bestehenden Definitionsversuchen [1, 12, 13], berücksichtigt aber insbesondere den Menschen in der Rolle des Entscheidungsträgers [14]. „Industrie 4.0"-Umgebungen synchronisieren dabei die digitale und physische Welt durch den Einsatz von Sensoren und Aktuatoren [14] und vernetzten zuvor isolierte Wissensinseln. Der Mensch soll dabei in der Rolle des Erfahrungsträgers, Entscheiders und Koordinators flexibel in die Arbeitswelt eingebunden bleiben [13]. Auch eine vollständig automatisierte und menschenleere Fabrik wäre zwar ein denkbares Zukunftsszenario, scheint aber unrealistisch [15]. Viel wahrscheinlicher hingegen wäre ein „Werkzeug-" oder „Hybrid-Szenario", in welchem der Mensch die Kontrolle über die Prozesse behält [15]. Dies erfordert aber auch, dass sich die Mitarbeiter das notwendige Wissen über die Produktionsabläufe aneignen können, trotz eines höheren Grades an Automatisierung. Kagermann [13] stellt in Aussicht, dass Mitarbeiter in diesen Szenarien eben nicht nur als „Maschinenbediener" fungieren sondern ein interdisziplinäres Wissen haben um adäquat auf die jeweilige Situation reagieren und entscheiden zu können. Somit ändern sich die Kompetenzanforderungen an den einzelnen Mitarbeiter grundlegend: Während in klassischen Produktionsumgebungen fokussiert Expertise in einem Spezialgebiet gefordert ist, zeichnen sich „Industrie 4.0"-Umgebungen durch einen breiten Kompetenzbedarf aus, der mehrere Fachdisziplinen vereint oder sogar ganz neue Berufsbilder hervorbringen soll [16].

2.1 Wissensmanagement während der Transition zu Industrie 4.0

Es stellt sich hierbei jedoch die Frage, wie eine Transformation in der Praxis erfolgen kann und was eine solche Umstellung hin zu flexibel einsetzbaren Mitarbeitern mit Breitenwissen für ein Unternehmen langfristig bedeutet.

Kompetenzaufbau erfolgt in konventionellen Unternehmen meist im Rahmen der Personalentwicklung (PE). Maßnahmen der PE unterstützen Mitarbeiter ihre Beschäftigungsfähigkeit in einem sich dynamisch veränderten Berufsumfeld zu erhalten [17]. Im Gegensatz zur Organisationsentwicklung zielt sie auf den einzelnen Menschen ab und fasst alle geeigneten und geplanten Anstrengungen zusammen, welche die individuelle berufliche Handlungskompetenz der Belegschaft entwickeln und erhalten [18]. Die PE folgt den Phasen (1) Bedarfsanalyse, (2) Konzeption, (3) Durchführung und (4) Transfer [19]. Während die Bedarfsanalyse und die Konzeption im Vorfeld einmal durchlaufen werden müssen, können die Durchführung und der Transfer wiederholt für jeden Mitarbeiter repetitiv stattfinden. Damit können im konventionellen Rahmen viele Mitarbeiter auf den gleichen Wissensstand gebracht werden. Durch die Flexibilisierung der Arbeit ist dies aber nicht mehr notwendig und muss stattdessen einem Prozess weichen, der das Lernen mit den tatsächlichen, individuellen Arbeitsprozessen synchronisiert [20]. Damit wird die Lernaktivität aber auch nicht mehr von außen initiiert, d.h. der Bedarf wird nicht zentral erkannt (Bedarfsanalyse), sondern entsteht ad-hoc anhand der aktuellen Situation.

Dies erfordert den Einsatz *didaktischer Datenbanken*, welche die Lerninhalte modularisiert abspeichern und verteilt zur Verfügung stellen [20]. Mobile Endgeräte bringen den Zugang zu diesen didaktischen Datenbanken nahe an den Arbeitsplatz [21] oder in beliebige andere Kontexte. Je nach Lernziel ist der Kontext für die Wissensvermittlung aber entscheidend. So ist beispielsweise das Wissen über die Wartung einer Maschine in einem Seminarraum anders, als die Situation des Lernenden, wenn er in einer Werkhalle direkt vor der Maschine steht. Es gibt vier Kontexte, in denen Lernen stattfinden kann [7]: (1) *Irrelevanter Kontext*: der Lerner hat keine wesentliche Beziehung zur aktuellen Umgebung. (2) *Formalisierter Kontext*: Ein Ort mit einer organisatorischen Funktion, welcher eine Gruppe von Lernern innerhalb desselben Kontextes (z.B. Klassenraum) synchronisiert. (3) *Physischer Kontext*: Der Ort, an dem sich der Lerner aufhält, ist für das Lernen relevant. (4) *Sozialisierender Kontext*: Lerner teilen dauerhafte, zwischenmenschliche Beziehungen. Im physischen und sozialisierenden Kontext kann Lernen arbeitsprozess-begleitend erfolgen. Der physische Kontext ist dabei besonders relevant, da er eine direkte Kontextualisierung des Wissens anbietet: Informationen sind direkt am Ort des Geschehens verfügbar; der Lerner kann reale Erfahrungen sammeln und sein gelerntes Wissen anwenden und erweitern. Lernen durch Erfahrungsepisoden (vgl. Experiential-Leanring [22]) gilt dabei als leistungsfähige Methode um kausale Zusammenhänge durch eigenes Erleben zu begreifen. Zwar rücken die Lerninhalte heute schon mit der Unterstützung durch mobile Technologien nahe an den Einsatzort heran, lassen sich aber trotzdem nicht vollständig darin integrieren. So bleiben z.B. die Transaktionskosten des einzelnen Mitarbeiters (suchen, konsumieren und anwenden) weiterhin bestehen. Die Frage wie diese Integration zu leisten ist bleibt unbeantwortet.

Einen Anhaltspunkt könnten die neuen Möglichkeiten geben, die im Rahmen einer „Industrie 4.0"-Vision technische Grundlagen schaffen könnten, um eine Umgebung zu erstellen die Lernkomponenten durch einen cyber-physischen Ansatz direkt in die Produktionsumgebung integriert. Dadurch werden die Produktionsmittel selbst Teil des Lernarrangements.

Die Erwartungen an diese technologische Entwicklung sind jedoch zum heutigen Zeitpunkt noch sehr spekulativ und reichen weit in die Zukunft. Visionen umfassen mehrere Stufen der Entwicklung von dem heutigen Bild der Fabrik als *Netzwerk mechatronischer Systeme* hin zu einer sich selbstoptimierenden, intelligent handelnden Produktionsumgebung [23]. Folgendes längeres Zitat zeigt wie weit diese Spekulationen gehen: *„Cyber-physische Systeme können ihre Umwelt unmittelbar mit ihrer entsprechenden Sensorik erfassen, sie mit Hilfe weltweit verfügbarer Daten und Dienste auswerten, speichern und sie können mit Hilfe von Aktoren auf die physikalische Welt einwirken. […] Diese CPS können sich dann vernetzen und autonom und dezentral – also ganz im Zeichen dieser selbstähnlichen Produktionsfraktale – Netzwerke aufbauen und sich eigenständig selbst optimieren. Sie können im Zusammenspiel mit dem Menschen eigenständig Probleme lösen."* [1, Seite 16] Was jedoch fehlt ist eine realistische Betrachtung der kurzfristigen Potentiale, die mit dem heutigen Stand der Technik erreicht werden können, um die Grundlage für ein integriertes, mobiles Lernen auf dem Weg zur intelligenten Fabrik zu bilden.

In anderen Feldern werden aber schon heute kontextsensitive Lernumgebungen [24] erprobt. Zum Beispiel in der Medizin-Ausbildung. Hier existieren Ansätze, um das Lernsystem mit in Dummies verbauten Sensoren zu verbinden, wodurch das Lernsystem direkt Feedback über die ausgeführten Handlungen am Körper bekommt [25]. Auch im Industrieumfeld sind erste Praxisberichte verfügbar, die von einem erfolgreichen Einsatz von mobile Learning in der Produktion berichten [26]. Auch wurden beispielsweise in einem kollaborativen Action-Design-Research (ADR) Projekt zwischen 15 Industrieunternehmen der Raumfahrt- und Autoindustrie und dem Production Technology Center (PTC) der Universität West in Schweden über 2 Jahre hinweg eLearning Kurse entwickelt. Das Kurs-Design bestand aus formalen Vorlesungen und verfügbaren eLearning Technologien am Arbeitsplatz, zu Hause oder mobil [27].

Ein konkretes Vorgehen zur Umgestaltung der Personalentwicklung aufgrund der neuen Anforderungen ist noch wenig unterstützt. Mobiles kontextsensitives Lernen scheint hier Potential zu haben diese Lücke zu schließen. Bisherige Lösungen bringen die Lernumgebungen zwar nah an den Arbeitskontext, eine vollständige Einbettung ist mit den bisherigen mobilen Technologien aber noch problematisch. Durch die Entwicklungen und Visionen von Industrie 4.0 eröffnen sich grundlegend neue Möglichkeiten für die Erfassung und Interpretation des Arbeitskontextes, den Handlungen von Menschen und den Zuständen von Maschinen. Vor diesem Hintergrund fragen wir nach dem Potential von mobilem Lernen unter den geänderten Voraussetzungen durch Industrie 4.0.

3 Methode

Die Forschungsfrage soll mittels Design-Science [28] beantwortet werden. Dabei versucht Design Science im Allgemeinen menschliche und organisatorische Fähigkeiten durch die Gestaltung neuer Artefakte zu verbessern [28]. Insbesondere dort, wo ein Routine-Design aufgrund von fehlendem Wissen oder aufgrund der hohen Komplexität der Situation und der darin notwendigen menschlichen Interaktionen nicht zielführend ist, liegt die Stärke dieses Ansatzes [28]. Die Projektstruktur orientiert sich dabei an der Design-Science-Research-Methodology [29], die sich in sechs Phasen gliedert: (1) Problemidentifikation, (2) Definition des Ziels, (3) Design und Entwicklung, (4) Demonstration, (5) Evaluation und (6) Kommunikation.

In dieser Publikation stehen die beiden ersten Phasen im Vordergrund, wodurch wir generische Probleme von konventionellen Lernarrangements in „Industrie 4.0"-Umgebungen aufzeigen und anhand einer Fallstudie an einem konkreten Beispiel veranschaulichen und detaillieren. Dennoch entwerfen wir erste, mögliche und zukunftstaugliche Lernarrangements für den diskutierten Fall.

Das Vorgehen ist kompatibel mit dem Evaluations-Framework von Sonnenberg und vom Brocke [30], die gezielte Evaluationsschritte nach jeder DSR-Aktivität des DSRM-Frameworks [29] vorsehen. In diesem Fall ist dies die Ex-Ante-Evaluation der Probleme. Konkrete und abstrakte Probleme sowie Lösungen halten wir mit Hilfe des Design-Theorizing Frameworks [31] auseinander. Wir beginnen also mit dem abstrakten Problem, welches durch die Fallstudie zum spezifischen Problem erweitert wird und geben anschließend einen ersten Ausblick auf die abstrakte Lösung. Das abstrakte Problem beschreiben wir (analog zum Vorgehen der Explanatory Design-Theory [32]) mit Hilfe von Requirements, die als Gegebenheiten der Problemsituation (Conditions) und notwendige Fähigkeiten (Capabilities) der Akteure beschrieben werden.

4 Abstrakte Problemsituation

Ausgehend von den im Literaturteil aufgezeigten Grundlagen beschreiben wir in diesem Abschnitt die generischen Anforderungen an zukünftige Lernarrangements im Industrieumfeld. Wie im Literaturteil bereits ausgeführt ist die menschliche Kapazität für Wissensaufnahme begrenzt **(CON1)**. Daraus folgt auch, dass nicht gleichzeitig Tiefenwissen in allen Bereichen erlangt werden kann [11] und daher bei Bedarf Wissen aufgebaut werden muss, sofern es noch nicht vorhanden ist. Eine weitere Gegebenheit entsteht durch die Vergänglichkeit von gelerntem Wissen **(CON2)**: Wissen, dass nicht aktiv benötigt wird, gerät in Vergessenheit [33]. Es sollte daher nur das Wissen vermittelt werden, dass auch zeitnah in der Arbeitssituation benötigt wird. „Just-in-Case"-Lernen muss daher eher einem „Just-in-Time"-Ansatz weichen [34]. In einem formellen Lernkontext ist dies schwierig zu realisieren, weil der momentane Wissensbedarf instabil ist und sich auch der Anwendungszeitpunkt des vermittelten Wissens nur unscharf voraussagen lässt.

Daher formulieren wir folgende **generische Anforderung (CAP1)**: *Der zukünftige Lerner erhält Grundlagenwissen individualisiert, modularisiert und zeitlich flexibel.*

Wie allgemein gefordert, muss Erfahrungswissen auch erhalten werden. Zu einem gewissen Grad lässt sich dies durch sozialisierende Ansätze lösen. Dieser Ansatz stößt jedoch schnell an seine Grenzen, da das vermittelte Wissen nur innerhalb einer (kleinen) Gruppe von involvierten Akteuren verbleibt.

Daher formulieren wir folgende **generische Anforderung (CAP2)**: *Der zukünftige Lerner kann sein persönliches Erfahrungswissen an andere Lernende weitergeben und sein eigenes Wissen um das Erfahrungswissen anderer erweitern.*

Insbesondere dadurch, dass Lernen und Arbeiten zu einer untrennbaren Einheit werden soll, müssen sich die Lernaktivitäten zu einem Großteil nahtlos in die Arbeitsprozesse einfügen. Dabei soll der Wissenstransfer direkt am Arbeitsplatz ermöglicht werden.

Daher formulieren wir folgende **generische Anforderung (CAP3)**: *Der zukünftige Lerner verknüpft (z.B. durch Erfahrungsepisoden) selbstständig reale Problemlösungssituationen mit kontextspezifischen Lerninhalten und bindet Experten bei Bedarf ein.*

5 Fallstudie eines Unternehmens im Wandel zu Industrie 4.0

> *„[...] ich brauche einen Mitarbeiter, der Know-How hat von der spanenden Bearbeitung bis zur […] prüftechnischen Ablaufprogrammierung. Und das ist natürlich ein enormes Spektrum an Fachkenntnis."* (Zitat Führungskraft QS)

5.1 Betrachtetes Unternehmen

Die Automobilzulieferer AG (AZ AG) arbeitet schon im Wertstrom-Design und hat dieses jetzt bis auf Teamleiterebene heruntergebrochen. Zum Zeitpunkt der Erhebung ist diese Transformation bereits erfolgt und die Belegschaft arbeitet in diesem neuen Muster fachübergreifend. Besonders gut können sich die Auswirkungen dieser Transformation an den Kompetenzanforderungen der Mitarbeiter der Qualitätssicherung (QS) aufzeigen lassen, da diese Tätigkeit schon bisher informationslastig war und viel Erfahrungswissen von hoch spezialisierten Mitarbeitern gebildet wurde. Im neuen wertstromorientierten Design werden diese Spezialisierungs-Silos aufgebrochen und in fachübergreifenden Tätigkeiten entlang des Wertstroms transformiert. Der Wissensaustausch des erlernten spezifischen Wissens zwischen den Mitarbeitern der Qualitätssicherung bekommt damit einen höheren Stellenwert. Zusätzlich haben auch die Produktionsmitarbeiter eine erhöhte Eigenverantwortung für die Qualitätssicherung übernommen und müssen daher von der QS geschult werden.

Das Problem, welches wir hier betrachten können, erstreckt sich über zwei Phasen: (1) die Reorganisationsphase, in der die bestehende Belegschaft neu qualifiziert werden muss und (2) die kontinuierliche Phase in der das Wissen langfristig erhalten bleiben muss.

Zu beiden Phasen sammeln wir mit Hilfe einer Vor-Ort-Erhebung Probleme und setzen diese mit der Literatur in Verbindung.

5.2 Datenerhebung

Die Datenerhebung für die Fallanalyse fand zwischen 05.2015 und 11.2015 statt. An insgesamt 5 Tagen konnten im Unternehmen Beobachtungen gemacht und Interviews geführt werden. Insgesamt wurden 3 Interviews mit Produktionsmitarbeitern, 5 Interviews mit Mitarbeitern der Qualitätssicherung und 3 weitere Interviews mit Kontaktpersonen der QS im Unternehmen geführt. Die Interviewten der Produktion haben eine Berufserfahrung zwischen 5 und 27 Jahren, die der QS zwischen 6 und 33 Jahren und die Kontaktpersonen der QS haben bis zu 20 Jahre Erfahrung. Die Interviews mit der QS wurden über alle Hierarchieebenen geführt, vom Leiter der QS, über die Fachleiter bis hin zu den Prüfern. Die Interviews dauerten 33 bis 85, im Mittel etwa 50 Minuten. Sie wurden aufgezeichnet (Ton) und anschließend transkribiert. Die konkreten Probleme wurden aus den Interviews extrahiert, in Form von Problemszenarien dargestellt [vgl. 35] und dem Unternehmen anschließend zur Validierung in Schriftform zur Verfügung gestellt. Die Teilnehmer wurden vom Unternehmen vorgeschlagen. Die Teilnahme erfolgte auf freiwilliger Basis. Alle Befragten haben einen „Informed Consent“ unterzeichnet in dem sie über die ethischen Grundsätze und ihre Rechte aufgeklärt wurden – dazu gehörte insbesondere, dass die Weiterverarbeitung der Daten ausschließlich anonymisiert oder aggregiert erfolgt.

5.3 Konkrete Probleme der aktuellen Lernsituation im Unternehmen

Der Kompetenzaufbau bei der AZ AG findet zum heutigen Zeitpunkt in drei Stufen statt. Für neue Mitarbeiter wird dies formell in individuellen Einarbeitungsplänen festgehalten, für bestehende Mitarbeiter erfolgt dies bedarfsorientiert. Für die verschiedenen Stufen kommen unterschiedliche Lernarrangements zum Einsatz. Für Faktenwissen werden nach wie vor Grundlagenschulungen angeboten. Zum Transfer von Expertenwissen wird den auszubildenden Mitarbeitern ein Mentor zur Seite gestellt. In der letzten Stufe lernen die Beteiligten situativ vor Ort mit Unterstützung von erfahrenen Mitarbeitern bei Bedarf.

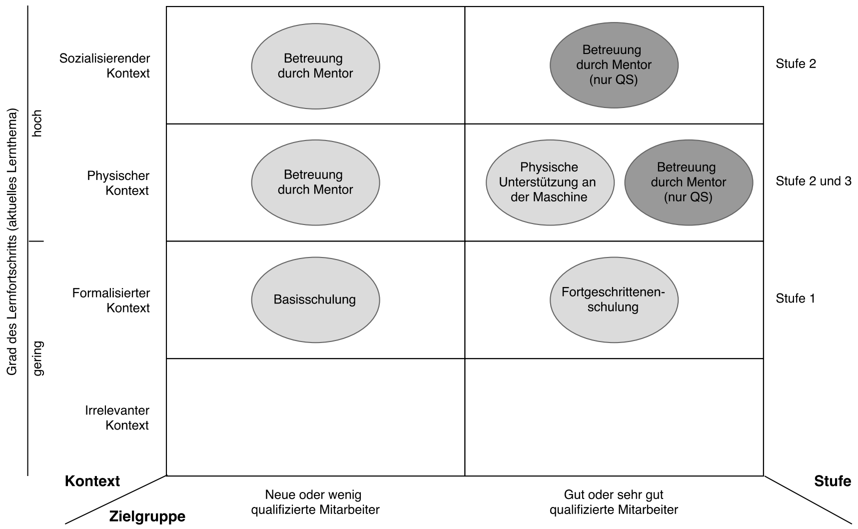


Abbildung 1. Aktuelle Lernarrangements bei der AZ AG

Viele Lernsituationen finden in der AZ AG mobil statt. Deshalb bietet es sich an, die Lernsituationen nach dem Mobile Learning Framework von Frohberg et al. [7] zu strukturieren, das Lernumgebungen in die vier Kontexte einteilt, welche bereits im Literaturteil beschrieben wurden. Da verschiedene Zielgruppen für die Arbeits- und Lernprozesse zentral sind, wird das Mobile Learning Framework um diese Dimension erweitert. Relevante Zielgruppen sind: wenig und gut bis sehr gut qualifizierte Mitarbeiter (diese können in Produktionsarbeiter auf dem Shopfloor und Mitarbeiter der Qualitätssicherung kategorisiert werden). Die Darstellung der Lernarrangements in Abbildung 1 zeigt den Zusammenhang zwischen Kontext und Zielgruppe, und der Zuordnung der Kontexte zu den drei Stufen.

Hierbei ist auch ersichtlich, für welchen Grad des Lernfortschritts des aktuellen Lernthemas der jeweilige Kontext eher geeignet ist (im physischen und sozialisierenden Kontext ist der Grund für das Lernen beispielsweise bereits vorhandenes Wissen anzuwenden und zu erweitern) [36]. Zur Erreichung der Qualifikation sind auf jeder Stufe Lernarrangements etabliert, welche im Folgenden auf ihre Probleme hin beschrieben werden.

5.3.1 Stufe 1: Mitarbeiterschulungen im formalisierten Kontext

Die aktuelle Praxis der Aus- und Weiterbildung bei der AZ AG basiert primär auf unterschiedlichen Arten von Schulungen und Workshops. Es gibt ein- bis zweitägige Basiskurse, die alle Mitarbeiter durchlaufen, um allgemeines Grundlagenwissen über Messmittel und deren Verwendung zu erhalten. Arbeitsplatzspezifische Wissensanforderungen werden über Einarbeitungspläne festgelegt:

„Im Einarbeitungsplan sind auch zum Beispiel Grundschulungen von allen Messmitteln, also Rundheitsmessgerät, Rauheitsmessgerät, Konturograph, Projektor, Optic Line […] enthalten […]“ (Zitat Führungskraft QS)

Weiterhin werden Schulungen für Fortgeschrittene angeboten, welche Prozesswissen (z.B. 5S, Kanban, Shopfloor-Management) vermitteln. Die Schulungen und Workshops werden sowohl für Produktionsmitarbeiter, als auch für Mitarbeiter der QS angeboten. Sie finden in einem formalisierten Kontext statt. Erfahrene Mitarbeiter der Qualitätssicherung bereiten diese Schulungen vor und planen die Termine zur Durchführung. Mitarbeiter der verschiedenen Abteilungen können dann an diesen Maßnahmen teilnehmen. Die Veranstaltungen finden in speziell dafür vorbereiteten Schulungsräumen und im Frontalunterricht statt. Dadurch entsteht das Problem (P1), dass die Mitarbeiter an einem bestimmten Zeitpunkt zu einem Thema geschult werden, ohne dass kontext- oder situationsspezifische Anforderungen an die Lehrinhalte existieren, was der generischen Anforderung (CAP1) des zeitlich flexiblen Wissensaufbaus entgegensteht:

„Sie (die Mitarbeiter, Anmerkung des Verfassers) haben es gelernt, [...] mit einem bestimmten Messmittel […] umzugehen, die Programme entsprechend einzustellen. Dann sind sie vielleicht [...] ein halbes Jahr in einem anderen Bereich eingesetzt […]. Dann kommt er nach einem halben Jahr wieder zurück und dann hat er natürlich diese [...] gelernten Fertigkeiten [...] durchaus vergessen [...].“ (Zitat Führungskraft QS)

Erschwerend kommt als Problem (P2) hinzu, dass es durch den flexiblen Einsatz der Mitarbeiter schwierig ist, abteilungsübergreifende Termine für gemeinsame Schulungen zu finden, was die generische Anforderung (CAP1) weiter substantiviert. Im Extremfall fallen die bereits organisierten Schulungen aufgrund von spontanen Änderungen im Arbeitsplan sogar aus:

„[…] wir bieten bestimmte Schulungen an, halten uns ein Zeitfenster frei [...]: Dann stehen wir bereit, Messmittel, alles haben wir entsprechend hergerichtet und dann kommt [...] kein Mitarbeiter, weil plötzlich krank, nicht da, muss anderweitig eingesetzt werden, weil wir einen Engpass an einer anderen Maschine haben.“ (Zitat Führungskraft QS)

5.3.2 Stufe 2: Mentoring im sozialisierenden Kontext

In der Qualitätssicherung der Fertigung teilen sich zwei erfahrene Mitarbeiter unterschiedlicher Spezialisierungen ein Büro, lösen auftretende Probleme gemeinsam direkt auf dem Shopfloor und übertragen so gegenseitig Wissen auf einander. D.h. beide Mitarbeiter laufen für mehrere Wochen zusammen an die Maschinen und Anlagen, um die Produktionsmitarbeiter zu unterstützen und Probleme zu lösen. Im

Falle von neuen oder weiter auszubildenden Mitarbeitern erfolgt das Mentoring nur in eine Richtung. Für diese Mitarbeiter wird der Mentor im Einarbeitungsplan festgelegt.

Das Problem (P3) dieses Ansatzes ist, dass sich der Lehrinhalt ausschließlich an den tatsächlich auftretenden Gegebenheiten und Problemen definiert. Das heißt, dass einerseits nur Wissen über Probleme transferiert wird, die in diesem Zeitraum auftreten. Andererseits verbleibt das Gelernte ausschließlich in dieser Dyade und ist für andere Mitarbeiter und Mentoring-Gruppen somit nicht zugänglich, was der Anforderung (CAP2) eines teilbaren Erfahrungswissens im Weg steht:

> *„Die Störung an sich wird zwar dokumentiert, aber der Weg dahin und die Störungsbeseitigung, das ist also wenig beschrieben, […]. Das ist [...] die Erfahrung des Mitarbeiters, […]" (Zitat Führungskraft QS)*

5.3.3 Stufe 3: Situative, persönliche Unterstützung im physischen Kontext

Hierbei tritt die QS als Helfer bei schwierigen Problemen und Störungen gegenüber den Produktionsmitarbeitern auf. Diese Experten sind eine begrenzte Ressource, da sie über Langzeiterfahrung im Unternehmen verfügen müssen:

> *„Die Mitarbeiter, die lange in dem Bereich tätig sind, die beheben die Störung, sagen dem Mitarbeiter: Hei, so oder so musst du es das nächste Mal machen, [...]" (Zitat Führungskraft QS)*

Die Situation wird noch verschärft, da die hilfesuchenden Mitarbeiter in der Regel keine Vorleistungen erbringen. D.h. bekommt der Experte vorher keine Information von seinen Kollegen vor Ort (z.B. vom Einsteller), muss er sich beim Eintreffen erst selbst ein Bild der Situation machen:

> *„[…] wenn er bloß sagt: Ja, der (Kolben, Anmerkung des Verfassers) geht nicht rein [...]: Ja, dann habe ich verschiedene Kriterien wie Durchmesser, Rundheit oder Konizität oder wie auch immer. Wenn ich da schon gewisse Sachen von vornherein als nicht so dringlich erachten kann, dann kann ich mich auf die [...] möglicherweise eigentliche Ursache besser konzentrieren." (Zitat QS-Mitarbeiter)*

Das Problem (P4) dieses Settings ist die Verfügbarkeit und Auslastung dieser Experten und bekräftigt die Anforderung (CAP3) nach selbstständigen Problemlösungskompetenzen. Fehlt diese Kompetenz, müssen die anfallenden Probleme warten, bis ein Experte verfügbar ist.

Zusammenfassend lässt sich also festhalten, dass in der kürzlich reorganisierten flexiblen Arbeitsumgebung Probleme entstehen, wenn Mitarbeiter (1) unabhängig vom Arbeitskontext geschult werden, (2) weiterhin in formellen Kontexten geschult werden, (3) sich Erfahrungswissen in Kleingruppen aneignen, dadurch Lernziele verpassen oder das eigene Wissen nicht weitergeben können und (4) durch Experten unterstützt werden, die ihnen die komplette Problemlösungsarbeit abnehmen.

6 Diskussion

Die identifizierten abstrakten Probleme lassen sich einerseits aus der Literatur argumentativ ableiten, mit dem Fallbeispiel konnten wir aber zeigen, wie sich diese in einem realen Unternehmen in einer Transitionsphase hin zu Industrie 4.0 manifestieren.

In den folgenden Sektionen bieten wir eine Diskussion an, wie sich die neuen Ansätze von Industrie 4.0 eignen würden, um eine angemessene Lernunterstützung zu ermöglichen.

6.1 „Industrie 4.0“-Technologien als Enabler zukünftiger Lernarrangements

Wir kombinieren die Ansätze des kontextsensitiven mobilen Lernens mit der Vision von cyber-physischen-Systemen, um die angeführten Anforderungen zu erfüllen. In einem Industrieumfeld lässt sich der Lernkontext durch den Zustand der Produktion und der aktuellen Tätigkeit der Mitarbeiter beschreiben. Soll ein Lernsystem kontextsensitiv sein, muss es Zugang zu diesen Informationen erhalten. In einem „Industrie 4.0“-Umfeld existieren diese Daten bereits in Form von vernetzten IoT-Systemen[1], die fortlaufend den Zustand der Maschinen überwachen. Dadurch werden Informationen zu (sich anbahnenden) Problemsituationen zugänglich, wodurch das Lernsystem die relevanten Inhalte offerieren kann, um damit die notwendige Unterstützung zur Problemlösung zu bieten und direkt ein problemorientiertes Lernen [vgl. 36] zu unterstützen.

Diese Sensoren erfassen (als Teil der Mensch-Maschine-Schnittstelle) auch die Handlungen der Bediener. Dies liefert dem Lernsystem Rückschlüsse auf die korrekte Ausführung der zur Problemlösung erforderlichen Operationen und der darin inhärent enthaltenen Lernziele. Dies transportiert somit den sensorbasierten Ansatz aus der Medizinausbildung [25] in die Produktionsumgebung.

Bei bislang unbekannten Problemen können Bedienungsabläufe aufgezeichnet werden. Damit wird Feedback über die Problemlösung geliefert (neues Wissen extrahiert), das anderen zur Verfügung gestellt werden kann.

Erkennt das System Lernmöglichkeiten und lässt die Umgebung eine Lernepisode zu, bietet das Lernsystem dies automatisch an. Dies unterstützt das individualisierte Lernen und kann den aktuellen Kontext berücksichtigen oder frei davon (irrelevanter Kontext) Weiterbildung ermöglichen. Tritt ein unbekanntes Problem auf, entscheidet das System, ob ein bekannter Lösungsweg vorhanden ist oder ein Experte herangezogen werden muss. Via Audio/Video-Telekooperation können die Experten bei Bedarf vorerst auch nur remote eingebunden werden. Wenn Bild und Ton aufgezeichnet wird, kann das so entstandene Material auch anderen Gruppen und Mitarbeitern zur Verfügung gestellt werden. Augmented Reality lässt die Grenzen zwischen Arbeits- und Lernkontext vollständig verschwinden. Dadurch, dass kein zusätzliches Gerät mehr in der Hand gehalten oder damit bedient werden muss, wird ein echt paralleles Arbeiten und Lernen möglich was die bisherigen Ansätze mobiler Lern-Technologien im Industrieumfeld erweitert [26].

[1] Internet of Things: „intelligente Objekte“, welche an den Cyberspace angebunden und miteinander vernetzt sind, unterstützen den Menschen, ohne abzulenken oder aufzufallen. [37]

6.2 Konzept zukünftiger Lernarrangements

Auf dieser Basis zielen die Gestaltungsüberlegungen daher auf das Konzept des mobilen, kontextsensitiven Lernens ab, um den zukünftigen Bedarf an Aus- und Weiterbildung umfänglich zu adressieren. Die schon in der Praxis erfolgreich angewendeten didaktischen Datenbanken [20] können durch den gezeigten integrierten Ansatz nochmals deutlich aufgewertet werden. Die damit für den Nutzer einhergehenden Transaktionskosten (Suchen und Konsumieren) könnten durch Automatisation unter Einbezug der Kontextinformation verringert werden.

Um die Auswirkungen dieser Entwicklungen zu zeigen, wollen wir nun kurz darstellen, wie sich die in der Fallstudie identifizierten Lernarrangements verändern ließen. Abbildung 2 zeigt eine Übersicht der Lernangebote nach einer solchen Transformation. Man sieht, dass der irrelevante Kontext jetzt nutzbar wird und große Teile der formellen Lehrsituationen aufnimmt. Entsprechend der Anforderung CAP1 werden Lerninhalte für Novizen und Fortgeschrittene dynamisch, orts- und zeitunabhängig abrufbar. Weiterhin können Qualifikationsmaßnahmen mit zugehörigen Lernkontrollen in diesen Kontext transportiert werden, die notwendig sind, um Arbeitsplatzanforderungen zu erfüllen und welche vorher ausschließlich dem formalisierten Kontext vorbehalten waren.

Die im „Industrie 4.0“-Umfeld mögliche Auswertung von Telemetriedaten für eine präventive oder Problem-Analyse ist nun ebenfalls orts- und zeitunabhängig möglich. Hierbei wird auch ein situatives Lernen [vgl. 38] unterstützt, indem anhand von Maschinendaten zeitnah die Auswirkungen von geänderten Einstellungen analysiert werden können.

Gänzlich auflösen lässt sich der formalisierte Kontext aber nicht [39] und würde auch weiterhin für Pflichtveranstaltungen, wie Grundlagenschulungen und praktische Übungen/Prüfungen im Rahmen von Qualifikationsmaßnahmen, genutzt werden können.

Im Zentrum der Innovation steht die Unterstützung des arbeitsprozessbegleitenden Lernens im physischen Kontext – d.h. direkt an der laufenden Maschine und parallel zur Produktion, was direkt die Anforderung CAP3 adressiert. Sowohl hierbei, als auch bei der, in komplexeren Problemsituationen notwendigen, Expertenunterstützung würde das System den Lernenden mit bekanntem Problemlösungswissen unterstützen und ermöglichen, neu erarbeitetes Wissen zu externalisieren und im Unternehmen zu verteilen um somit die Anforderung CAP2 zu unterstützen. Dabei soll der soziale Kontakt zwischen Ratsuchendem und Experten ungestört verlaufen können – das System dient hier lediglich zur (Nach-) Dokumentation und zum Wissenstransport.

Eine weitere Innovation stellt die Unterstützung der Belegschaft im sozialisierten Kontext durch eine entsprechend gestaltete Community für die Diskussion und Lösung von nicht-akuten Problemen (akute Probleme müssen sofort gelöst werden) dar. Hier findet ein Lernen direkt zwischen den Mitarbeitern auf Shopfloor-Ebene statt.

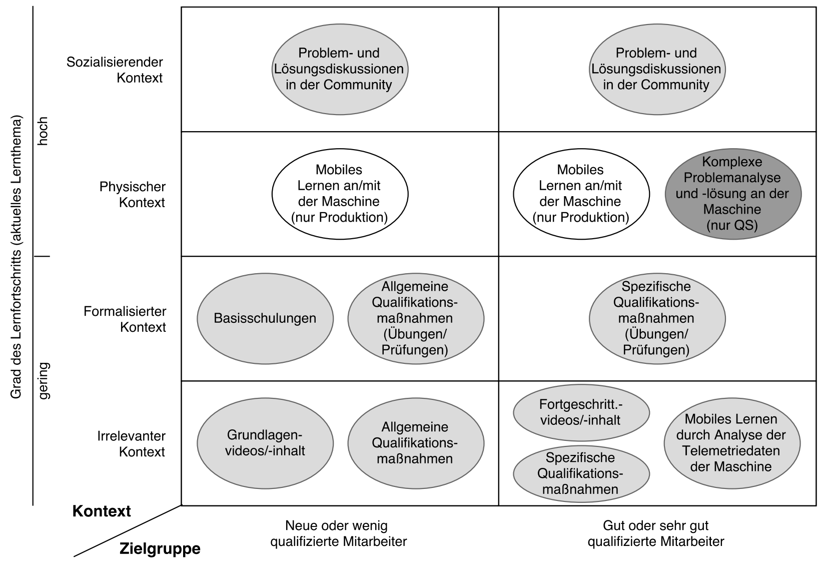


Abbildung 2. Lernarrangements bei der AZ AG mit optimierter Lösung

Die ohnehin schon existierenden sozialisierten Kleingruppen an den Anlagen, welche aus den Mitarbeitern der entsprechenden Arbeitsschichten bestehen, werden somit ausgedehnt und schichtübergreifend, anlagenübergreifend, zeit- und ortsunabhängig in ihrer Kommunikation unterstützt. Durch die individuelle Förderung der Mitarbeiter werden Kompetenzen anhand des tatsächlichen Bedarfs aufgebaut. Bedarfsanalysen im klassischen Sinne entfallen, da sich das System selbst reguliert – nicht mehr gebrauchtes Wissen wird vergessen und benötigtes Wissen punktgenau akquiriert. Die Personalkompetenz passt sich damit auch global den wechselnden Bedürfnissen des Unternehmens an. So behält der Mensch seine wichtige Rolle als Entscheidungsträger auch bei weiterer Zunahme der Komplexität in der Produktion.

Dieses Paper beschreibt dabei erst den Anfang des Design Research Prozesses. Momentan ist die Erkenntnis über mögliche Lösungen noch limitiert. Einerseits, weil sich die Untersuchung bisher nur auf einen konkreten Fall stützt und, weil die vorgeschlagenen Lösungen noch nicht im Unternehmen eingeführt sind. Andererseits zeigt sich durch diese Konstellation die praktische Relevanz am Fallbeispiel und im Laufe des Projektes können die vorgeschlagenen Interventionen in einem echten Kontext erprobt werden.

Acknowledgements

Dieses Projekt wird finanziert durch das Forschungs- und Innovationsprogramm Horizont 2020 der Europäischen Union unter der Projektnummer 636778.
Wir danken Alexander Richter für seine Unterstützung.

Referenzen

1. Bauernhansl, I.T.: Die Vierte Industrielle Revolution–Der Weg in ein wertschaffendes Produktionsparadigma. In: Industrie 4.0 in Produktion, Automatisierung und Logistik. pp. 5–35. Springer (2014).
2. Stocker, A., Brandl, P., Michalczuk, R., Rosenberger, M.: Mensch-zentrierte IKT-Lösungen in einer Smart Factory. e & i Elektrotechnik und Informationstechnik. 131, 207–211 (2014).
3. Davis, J., Edgar, T., Porter, J., Bernaden, J., Sarli, M.: Smart manufacturing, manufacturing intelligence and demand-dynamic performance. Computers & Chemical Engineering. 47, 145–156 (2012).
4. Hermann, M., Pentek, T., Otto, B.: Design Principles for Industrie 4.0 Scenarios. In: 2016 49th Hawaii International Conference on System Sciences (HICSS). pp. 3928–3937. IEEE (2016).
5. Nonaka, I., Takeuchi, H.: The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press (1995).
6. Polanyi, M.: The tacit dimension. Routledge & Kegan Paul, London (1966).
7. Frohberg, D., Göth, C., Schwabe, G.: Mobile learning projects–a critical analysis of the state of the art. Journal of computer assisted learning. 25, 307–331 (2009).
8. Traxler, J.: Current state of mobile learning. In: Mobile learning. pp. 9–24. AU Press, Athabasca University, Edmonton (2009).
9. Kearney, M., Schuck, S., Burden, K., Aubusson, P.: Viewing mobile learning from a pedagogical perspective. Research in learning technology. 20, (2012).
10. Hou, H.-T., Wu, S.-Y., Lin, P.-C., Sung, Y.-T., Lin, J.-W., Chang, K.-E.: A Blended Mobile Learning Environment for Museum Learning. Educational Technology & Society. 17, 207–218 (2014).
11. Grant, R.M.: Toward a knowledge-based theory of the firm. Strategic management journal. 17, 109–122 (1996).
12. Scheer, P.D.D. h c mult A.-W.: Industrie 4.0: Von der Vision zur Implementierung. In: Obermaier, R. (ed.) Industrie 4.0 als unternehmerische Gestaltungsaufgabe. pp. 35–52. Springer Fachmedien Wiesbaden (2016).
13. Kagermann, H.: Chancen von Industrie 4.0 nutzen. In: Industrie 4.0 in Produktion, Automatisierung und Logistik. pp. 603–614. Springer (2014).
14. Nelles, J., Kuz, S., Mertens, A., Schlick, C.M.: Human-centered design of assistance systems for production planning and control: The role of the human in Industry 4.0. In: 2016 IEEE International Conference on Industrial Technology (ICIT). pp. 2099–2104. IEEE (2016).
15. Ahrens, D., Spöttl, G.: Industrie 4.0 und Herausforderungen für die Qualifizierung von Fachkräften. In: Digitalisierung industrieller Arbeit. pp. 184–205. Nomos Verlagsgesellschaft mbH & Co. KG (2015).
16. Stich, V., Gudergan, G., Senderek, R.: Arbeiten und Lernen in der digitalisierten Welt. In: Digitalisierung industrieller Arbeit. pp. 108–131. Nomos Verlagsgesellschaft (2015).
17. Schermuly, C.C., Schröder, T., Nachtwei, J., Kauffeld, S., Gläs, K.: Die Zukunft der Personalentwicklung. Zeitschrift Für Arbeits-Und Organisationspsychologie A&O. (2012).
18. Kauffeld, S.: Nachhaltige Weiterbildung. Betriebliche Seminare und Trainings entwickeln, Erfolge messen, Transfer sichern. Springer-Verlag Berlin Heidelberg, Berlin (2010).

19. Krisor, S.M., Rowold, J., Block, C.: Personalentwicklung. In: Human Resource Management. pp. 173–185. Springer (2015).
20. Negri, C.: Angewandte Psychologie für die Personalentwicklung. Springer (2010).
21. Engert, V.: Mobile Lernmöglichkeiten in der Automobilindustrie. In: de Witt, C. and Sieber, A. (eds.) Mobile Learning. pp. 205–217. Springer Fachmedien Wiesbaden (2013).
22. Kolb, D.A.: Experiential learning: Experience as the source of learning and development. Prentice Hall, Englewood Cliffs, NJ (1984).
23. Schlick, J., Stephan, P., Loskyll, M., Lappe, D.: Industrie 4.0 in der praktischen Anwendung. In: Industrie 4.0 in Produktion, Automatisierung und Logistik. pp. 57–84. Springer (2014).
24. Huang, Y.-M., Chiu, P.-S.: The effectiveness of a meaningful learning-based evaluation model for context-aware mobile learning. British Journal of Educational Technology. 46, 437–447 (2015).
25. Wu, P.-H., Hwang, G.-J., Su, L.-H., Huang, Y.-M., others: A Context-Aware Mobile Learning System for Supporting Cognitive Apprenticeships in Nursing Skills Training. Educational Technology & Society. 15, 223–236 (2012).
26. deWitt, C., Gloerfeld, C.: Mobile Learning - Process oriented learning and informing in changing working environments. Coordination, scientific guidance and didactical consulting, social network for Mobile Learning in working processes. (2013).
27. Hattinger, M., Eriksson, K.: Action Design Research: Design of e-WIL for the Manufacturing Industry. In: AMCIS 2015, Puerto Rico. pp. 1–14 (2015).
28. Hevner, A., March, S., Park, J., Ram, S.: Design Science in Information Systems Research. Management Information Systems Quarterly. 28, 75–105 (2004).
29. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. JMIS. 24, 45–77 (2007).
30. Sonnenberg, C., vom Brocke, J.: Evaluations in the science of the artificial–reconsidering the build-evaluate pattern in design science research. In: International Conference on Design Science Research in Information Systems. pp. 381–397. Springer (2012).
31. Lee, J.S., Pries-Heje, J., Baskerville, R.: Theorizing in design science research. In: Service-Oriented Perspectives in Design Science Research. pp. 1–16. Springer (2011).
32. Baskerville, R., Pries-Heje, J.: Explanatory design theory. Business & Information Systems Engineering. 2, 271–282 (2010).
33. Loftus, G.R.: Evaluating forgetting curves. Journal of Experimental Psychology: Learning, Memory, and Cognition. 11, 397 (1985).
34. Pimmer, C., Pachler, N.: Mobile learning in the workplace: Unlocking the value of mobile technology for work-based education. Increasing Access. 193 (2014).
35. Rosson, M.B., Carroll, J.M.: Usability engineering: scenario-based development of human-computer interaction. Morgan Kaufmann (2002).
36. Haake, J., Schwabe, G., Wessner, M.: CSCL-Kompendium 2.0: Lehr-und Handbuch zum computerunterstützten kooperativen Lernen. Walter de Gruyter (2012).
37. Mattern, F., Floerkemeier, C.: Vom Internet der Computer zum Internet der Dinge. Informatik-Spektrum. 33, 107–121 (2010).
38. Brown, J.S., Collins, A., Duguid, P.: Situated cognition and the culture of learning. Educational researcher. 18, 32–42 (1989).
39. Sharples, M., Taylor, J., Vavoula, G.: A theory of learning for the mobile age. In: Medienbildung in neuen Kulturräumen. pp. 87–99. Springer (2010).

Yes You Can – Empowering Lecturers to Simulate Collaboration among Learners in the Disciplines of Problem-Solving and Critical Thinking Regardless of Class Size

Sarah Oeste-Reiß[1], Eva Bittner[2], Matthias Söllner[1,3]

[1] University of Kassel, Information Systems, Kassel, Germany
{oeste-reiss,soellner}@uni-kassel.de
[2] University of Hamburg, Information Systems, Hamburg, Germany
{bittner}@informatik.uni-hamburg.de
[2] University of St.Gallen, Institute of Information Management, St. Gallen, Switzerland
{matthias.soellner}@unisg.ch

Abstract. Fostering higher-level learning in the disciplines of problem-solving and critical thinking becomes important when educating knowledge workers. By taking part in collaborative learning (CL) activities, e.g., interactive discussions, learners have the chance to develop, defend, and critique positions. However, implementing CL activities is often complex because this requires knowledge in designing effective collaboration. We build on insights from learning and collaboration engineering literature to develop an IT-based Collaborative-Learning-Pattern Approach (CLPA) that consists of two patterns, each describing a process design – one for training problem-solving, and the other for attaining critical thinking abilities. To evaluate the CLPA, we use simulations, walk-throughs among lecturers, and pilot-tests among students. Results show that the CLPA empowers lecturers to implement respective activities in the classroom, takes into account pedagogical demands, and satisfies lecturers as well as learners. We contribute several findings toward a design theory for empowering lecturers to implement CL activities in their classes.

Keywords: collaborative learning, collaborative-learning-pattern approach, higher-level learning, collaboration engineering, design science research.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Oeste-Reiß, S.; Bittner, E.; Söllner, M. (2017): Yes You Can - Empowering Lecturers to Simulate Collaboration among Learners in the Disciplines of Problem-Solving and Critical Thinking Regardless of Class Size, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 761-775

1 Introduction

Approaches for training *higher-level learning (HLL)* on the upper levels of Bloom's taxonomy (apply, analyze, evaluate, create) [1] in the disciplines of problem-solving and critical thinking are becoming increasingly important in the digital age, which is characterized by an increasing availability of information. Furthermore, to remain competitive, innovative companies are in demand for highly educated knowledge workers. Competences such as teamwork and communication abilities are highly relevant as well [2]. The performance of knowledge workers depends on the degree to which they master those skills. Thus, universities have to provide learning experiences that help learners to develop those skills. However, traditional lectures – characterized by a low level of interaction among learners and a focus on factual knowledge [3] – are still popular. The reasons are for example declining state funding and increasing student numbers [4]. This means that learners often lack the chance to develop, defend, and critique positions, which would be vital for achieving HLL. Collaborative learning (CL) approaches ground on insights from *constructive learning theory* that posits that learning occurs by experiencing an environment through interactions with other individuals [5]. These approaches seem to be promising when it comes to overcoming existing shortcomings. However, CL approaches that focus on HLL are typically less predictive and hardly replicable, demand an understanding of how to design effective collaboration, and do not restrict learners in their experiences [6]. Lecturers lack validated out-off-the-box techniques to conduct and stimulate CL activities among learners. While lecturers struggle with less predictive and hardly replicable learner interactions and outcomes, learners struggle with CL techniques in terms of HLL tasks. These tasks provide learners with a problem situation. Such situations require that learners develop a solution that represents a sophisticated understanding of knowledge concepts and their relationships and thus, train problem-solving abilities. Furthermore, these situations require that learners analyze and evaluate the situation and, thus, train critical-thinking abilities. Inexperienced learners not familiar with these HLL learning techniques are often overstrained since e.g. tasks seem to be unclear and open-ended; instructions focus on learning content, but often do not provide training or guidance on how to proceed through the CL experience for HLL. In contrast to constructivist learning literature that argues learning processes should be ad hoc [6, 7], collaboration literature shows that process structures can under certain conditions increase the number, quality, and creativity of ideas a group creates. They may also increase the number of communication cues exchanged within a group, and improve the quality of its work products while reducing cognitive load [8]. Most individuals – lecturers as well as learners – do not have an intuitive grasp of effective collaboration. In cases of inventing ad hoc collaboration, most groups tend to be ineffective [8]. This leads to the assumption that CL experiences may benefit from systematically designed collaboration that guides lecturers and learners. Therefore, applying insights from collaboration literature to the domain of learning might be a solution. A design methodology is needed that a) provides procedural guidance on how to split structure and that describes CL activities for HLL in a way that helps lecturers and learners proceed through CL activities in a predictive and effective way; and b) helps lecturers implement CL activities for HLL as building blocks in their classes. In that context, *collaboration engineering* is an approach that designs and deploys high-value recurring tasks and transfers them to practitioners (lecturers, learners) without the ongoing support from expert facilitators [8].

The goal of this paper is to help lecturers and learners overcome this challenge by answering the following research question: *How can CL knowledge be packaged in a reusable way so that it comprises sufficient collaboration techniques to empower lecturers (and learners) to conduct (and follow) CL activities for HLL in the classroom?* The objective of this paper is to develop the *Collaborative-Learning-Pattern Approach (CLPA)* comprising two process designs inherent in patterns for enhancing HLL – the *Problem-Solving Pattern (PSP);* and the *Critical-Thinking Pattern (CTP).* The *design goals* of CLPA are: (1) to help lecturers enhance CL activities for HLL in the areas of problem-solving and critical thinking in classes in a predictive way; (2) to help learners proceed through CL activities with assisting guidance on collaboration. We focus on these two patterns for two reasons. First, they enhance cognitive processes that refer to applying, analyzing, evaluating, and creating knowledge, and thus focus on the upper levels of Bloom's taxonomy [1]; and second, they help enhance skills relevant for knowledge workers such as teamwork and communication. Each pattern represents a design for a reusable and structured collaboration process that packages sufficient collaboration expertise so that non-experts (lecturers, learners) can execute and follow a well-designed work practice without training in tools and techniques. We follow the idea of patterns, because *patterns* "[…] exist as means of deriving useful solutions to recurring problems within specific contexts" [9]. Consequently, a pattern describes a recurring problem as well as the core of the solution for that problem in such a way that the solution can be used unlimitedly [9]. To guide our design choices, we rely on insights from collaboration engineering literature and use the *six-layer-model* to design and present the CLPA with its patterns [10]. The layers comprise the definition and configuration of a group *goal*, *products*, *activities*, *procedures*, *tools*, and *behavior* [10].

2 Design Science Research Framework

In this paper, we report a Design Science Research (DSR) study and structure our paper along Hevner's 2007 [11] three-cycle view (Figure 1). First, we start the *relevance cycle* by identifying a set of unsolved problems inherent in packaging sufficient collaboration expertise to enhance CL activities for HLL in the classroom (activity #1 | section 1). Second, we initiate the *rigor cycle* by drawing on justificatory knowledge from CL literature with respect to training problem-solving and critical thinking abilities (activity #2 | section 3). Thirdly, we start the *design cycle* and provide principles of form and function inherent in generalizable requirements for CL activities for HLL, and the CLPA design with its two patterns as a generalizable solution (activity #3 | section 4). In section 5, we complete several iterative design and relevance cycles by describing the procedures of testing three iterative exemplar instances of the CLPA in terms of a multi-method evaluation (activity #4 – simulation with designers │ activity #5 – walk-through with lecturers │ activity #6 – pilot-test with students in the classroom). The results show that the designed artifact of CLPA meets the design goals. In section 6, we complete the rigor cycle by adding *prescriptive knowledge*[1] [12] to the literature before we close with an outlook on future research in section 7. According to Gregor's 2006 descriptions [13], our CLPA resembles a theory of '*design and action*'. More precisely, it is of the type '*improvement*'. Lecturers can use CLPA to create their own instances [12].

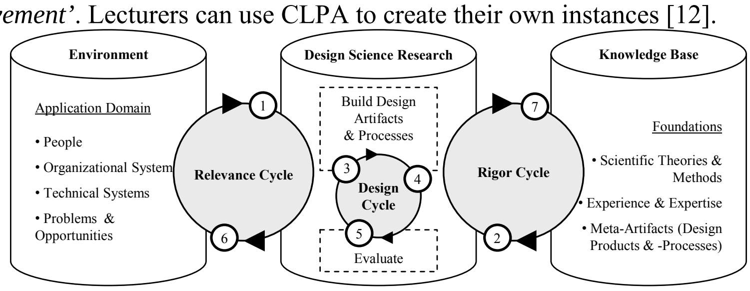


Figure 1. DSR Three-Cycle View in the Context of the Study

3 Theoretical Background of Collaborative Learning

Collaborative learning (CL) is based on constructivist learning theory [14]. According to this theory, learners learn from experiences that they gain through interactions with their environment and each other [15]. If well designed, CL may relieve the lecturer from some labor-intensive tasks, particularly in large classes, such as giving individual feedback on assignments. Learners benefit from such interactions in several ways: e.g. discussions can enable a direct exchange between learners that fosters reflection of knowledge, and thus, critical thinking; and can increase motivation, participation [16], and learning success [17]. This helps learners improve job-relevant competences like teamwork and communication [14]. The range of CL activities comprises discussions, co-construction of solutions, or giving mutual feedback. Literature on peer discussion of multiple-choice tasks, for example, describes positive learning effects when learners first reflect knowledge on their own, then discuss their choice with others, and finally re-evaluate their choice [18]. The co-construction of a solution, for example, helps learners explain and challenge ideas to each other, and stimulates knowledge creation [19]. Moreover, mutual assessment among learners has the potential to correct mistakes and to clarify unclear issues [20].

To enhance such CL experiences, lecturers need to respect several aspects. They have to ensure reciprocity in social interactions among learners, e.g. when it comes to direct feedback [21]. They also have to ensure that learners are responsible for their outcome [19] and that assignments and instructions are clear [22]. In a class, however, there are high- and less-experienced learners. Hence, it is hard for a lecturer to create a learning experience that challenges the top learners without losing the bottom

[1] *Prescriptive knowledge* describes artifacts designed by humans to improve the natural world. It is inherent in the form of models, methods, instantiations, and design theories.

learners. A shared understanding of knowledge concepts therefore is necessary to foster social interactions toward a development, modification, and reinforcement of shared mental models [23]. Van den Bossche et al. [24] identify *team learning behaviors* as follows: learners should express and share their individual understanding, and listen to each other (construction), discuss and clarify their understanding to reach mutual understanding (co-construction), and negotiate an agreement on a mutually shared perspective (constructive conflict). *Problem-based learning* is known to generate HLL. This is focused experiential learning that is organized around the investigation, explanation, and resolution of meaningful problems [25]. This way it refers to metacognitive knowledge on the upper levels of Bloom's taxonomy [1]. Learners collaborate in small groups and solve a problem. Depending on the assignment, the learners train *problem-solving abilities* by creating a common solution for a complex situation. They can also train *critical thinking abilities* by evaluating, analyzing, interpreting or explaining a problem situation with the aim of making a reflective judgement [26]. The lecturer facilitates and guides learners through the learning experience [27]. Although a wide range of variations of problem-based learning have evolved in literature and educational practice, there are some core characteristics: (1) learning needs to be learner-centered; (2) learning has to occur in small groups under the guidance of a tutor; (3) the tutor needs to act as a facilitator; (4) authentic problems are primarily encountered in the learning sequence, before a preparation has occurred; (5) the problems encountered are used as a tool to achieve knowledge and skills that are necessary for problem-solving [25]. Fifty years after problem-based learning had evolved; it was applied to various educational contexts. Much evidence suggests that it is more effective than traditional methods with regard to enhancing learners' problem-solving and critical thinking abilities. However, skeptics argue that it is ineffective because it provides only minimum guidance and therefore is too complex and not compatible with human cognitive architecture [28]. From a meta-study, Hmelo-Silver 2004 [27] derives a research agenda that calls for more work in the areas of collaboration, scaffolding structures for inexperienced learners, and approaches to overcome the lack of skilled facilitators. "Classrooms have more students than one person can easily facilitate, and learning to facilitate well is a challenge" [27]. She suggests techniques such as procedural facilitation or scripted cooperation to address this challenge.

4 The Collaborative-Learning-Pattern Approach (CLPA)

4.1 Generalizable Requirements to Enhance Collaborative Learning for HLL

In the following, we describe the design cycle of our study. We present generalizable requirements for enhancing CL activities for HLL (section 4.1) and then describe the CLPA with its two patterns as a generalizable solution (section 4.2). Following the DSR paradigm, we derive generalizable requirements (see Table 1) to design CL activities for HLL by completing a relevance cycle (section 1) and a rigor cycle (section 3). Based on this we derive the CLPA as our generalizable solution.

Table 1. Generalizable Requirements to Enhance CL Activities for HLL

Relevance cylce – lecturers' requirements (see the specific challenges/sources in brackets):
(GR 1) Set-up Guidance*:* The CLPA shall provide instructions for the task choice, as well as definition, set-up, and configuration of CL activities (lack of CL design experience of lecturers).
(GR 2) Facilitation Guidance: The CLPA shall provide detailed instructions on the facilitation actions, e.g. statements and questions the lecturer needs to work with during the CL experience (unpredictable moderation of CL).
Relevance cylce – learners' requirements (see the specific challenges/sources in brackets):
(GR 3) Simplified Process Structure*:* The CLPA shall divide CL into activities with defined subtasks (learners' resistance to open-ended and highly complex task structures).
(GR 4) Collaborative Interaction Support*:* The CLPA shall provide instructions on how interactions among learners should be organized in each phase (high cognitive load because of inventing ad-hoc collaboration parallel to task solving)
(GR 5) Clear Goal/Outcome Specifications*:* The CLPA shall define clear final and intermediate goals and outcomes for the learners for a specific task (risk to self-efficacy and satisfaction in case of transparency)
Rigor cycle – collaborative learning literature:
(GR 6) Individual Reflection*:* The CLPA shall support individual construction and reflection of knowledge [24].
(GR 7) Mutual Feedback*:* The CLPA shall provide structured support for constructive feedback, sense making [20, 24].
(GR 8) Consolidation of Solutions*:* The CLPA shall provide structured support for negotiating and consolidating different perspectives towards a shared solution [22, 24].
(GR 9) Access to Solution: Exemplary solutions shall be provided to all learners or discussed after the task completion (given the partly unpredictable outcome of CL, all learners shall have the chance to receive a correct solution) [19].
(GR 10) Task Responsibility in Small Breakout Groups*:* The CLPA shall assign distinct, complementary subtasks to breakout groups small enough for each learner to feel responsible for the result [19, 21].

4.2 The Pattern-Based Process Designs as a Generalizable Solution

The aim of the CLPA is to initiate predictive small-group CL activities for HLL in the disciplines of problem-solving and critical thinking. Thus, the CLPA comprises two patterns – the PSP (see section 4.2.1 [Table 3]), and the CTP (see section 4.2.2 [Table 4]). To develop and describe the CLPA we use the six-layer model [10] as a design methodology in order to apply insights from collaboration engineering literature, such as process restrictions and structuration of collaboration, to the domain of CL [10]. By following the layers we systematically derive a reusable process design for each pattern that structures CL in a sequence of activities with several outcomes. Our generalizable requirements guide our design choices. To conduct the two patterns of CLPA in a scalable manner, the lecturer has to prepare some conditions (see Table 2):

Table 2. General Conditions to Conduct the Patterns of CLPA

Parameter	Description
Problem situation	Define an overall complex problem situation with action items in which the subtasks become embedded. A problem situation is a situation that covers the intended content to be learned as well as the specific and unique contextual factors to be considered, and that considers the conceptual connections of the problem within the curriculum [29].
Choose and create task structure	Define 2 up to 15 independent subtasks that refer to learning objectives (task specifics described in each pattern); pay attention that its execution takes place in parallel sub-/ breakout groups.
Specify deliverables	Realize learning objectives within the demands of the group deliverable (e.g. visualization and explanation) and pay attention to the fact that it is easy to present in the plenary group.
Breakdown group structure	The whole class is the plenary group. A plenary group can be divided into at least 2 up to 15 subgroups (4 to 30 participants each), working simultaneously. A subgroup can be divided into several breakout groups (2 to 6 participants each) [30].
Dependencies groups and tasks	Each participant is part of a breakout group and works on a specific subtask (number of subtasks = number of breakout groups). A subgroup receives all subtasks.
Scalability	The problem situation and its subtasks can be assigned to more than one subgroup and their breakout groups. Use tools that provide a shared working space for all breakout groups.

4.2.1 Problem-Solving Pattern (PSP)

The *group goal* of PSP is that learners simultaneously and collaboratively clarify, discuss, and develop a solution for a subtask within two hours. To keep learners motivated, the task should be appealing to them, e.g. by being relevant for learners' future career or addressing their personal interests [29]. The collaboration helps learners to satisfy individual goals such as becoming qualified knowledge workers by experiencing HLL in the discipline of problem-solving as well as training teamwork and communication abilities. To operationalize the goal we use an instrumental *group product*: each subtask solution has to be reported as a group deliverable in the form of text and visualizations to illustrate all relevant knowledge concepts and their relationships in a correct and abstract manner and thus, new knowledge is created. To operationalize the group product we define group activities to structure the collaboration. The PSP comprises three distinct steps, each using a *thinkLet*[2] to structure *group activities*. While the learners work in a subgroup in step 1, they collaborate within breakout groups in steps 2 and 3. In step 1, each learner receives access to all subtasks. On their own, learners brainstorm solution ideas while having the chance to read the contributions from their teammates. This activates chunking and thus, cognitive mechanisms to build relationships among knowledge frames. Reading ideas from other learners triggers cognitive effects among the less-experienced learners. In step 2, learners are assigned to breakout groups, each of which receives a subtask with the deliverables from step 1. In the breakout groups, learners discuss, organize, and summarize contributions and add missing knowledge aspects. This helps them consider and juxtapose the knowledge to create a solution for a problem situation. In step 3, learners report the solution by using text descriptions and visualizations. The *tool* support provides shared writing pages (e.g., GSS with separated groups, text editing, visualizations [e.g., ThinkTank, GoogleDocs, Google Slides]; flip charts; cards) so that the learners are able to make contributions while reading the contributions from other learners (step 1), and to discuss with other learners and visualize their solution (steps 2, 3). After each step the tools generate a report of the group deliverables (e.g. list of ideas). The *group behavior* restricts learner interactions toward solving the task. After each step, learners are stopped from editing documents and become automatically assigned to their group (plenary group, subgroup, or breakout group). Learners receive guidance via clear instructions, enabling them to cope with subtasks, showing them how to complete the activities, giving them orientation, e.g. with a list of teammates. Table 3 illustrates the PSP and serves as a moderation plan.

Table 3. Problem-Solving Pattern (PSP): Overview and Moderation Plan

Learning objective	Apply, analyze, evaluate, create	**Task specification**	Content, context, connection
Individual goal	Training HLL in the discipline of problem-solving and, teamwork and, communication abilities.		
Group product	For each subtask a solution in the form of meaningful text and visualizations (e.g., storyline/scenes, slide show).		
Group changes	2: subgroup to breakout group; breakout group to plenary group		
Tool support	GSS – functionalities for separate groups, text editing, visualizations: e.g., ThinkTank, GoogleDocs, flip charts, cards		

[2] *thinkLets* are packaged collaborative activities that serve as validated building blocks [33].

	Group activity & general description	Group product & quality indicator	Group procedures (thinkLet, pattern of collaboration) [8]
Step 1: 20 min	Subgroups: Each subgroup receives all subtasks. Each learner brainstorms ideas to create a solution for a problem situation among all subtasks.	Product: Per subgroup, a document with a set of solution ideas for each subtask. Quality: Contributions are solution ideas that represent knowledge concepts in the form of meaningful keywords.	LeafHopper (brainstorm): For each subgroup a bundle of shared writing pages, each with a subtask. a) Explain learners the subtasks and how to contribute. b) Explain expectations regarding quality aspects of contributions. c) Prompt learners to work on subtasks in which they have the most expertise; to look at each subtask, read it, and add contributions. d) Indicate that learners will not be able to work on every subtask during the available time.
Step 2: 40 min	Breakout groups: Subgroups are divided into breakout groups, each assigned to a subtask. Learners discuss, organize, and summarize solution aspects for a subtask.	Product: Per breakout group, a document with a clarified and summarized set of solution aspects for each subtask. Quality: Organized, corrected, and completed solution aspects.	PopcornSort (organize): For each breakout group a shared writing page with a subtask. a) Explain and verify instructions and converge categories (not relevant, correct, missing aspects). b) Explain that contributions are to be assigned to categories? c) Summarize the correct aspects in a meaningful explanation.
Step 3: 60 min	Breakout groups: Learners report their solution in the form of text and visualizations. Plenary group: Lecturer and learners discuss exemplary solutions.	Product: Per breakout group a report for a subtask in the form of text and visualizations. Quality: Report comprises all relevant knowledge concepts in the form of text and meaningful visualizations.	BucketBriefing (clarify): For each breakout group a shared writing page for text and visualization. a) Explain that learners are to work on the shared writing page. b) Explain to learners quality criteria to report the solution. c) Explain that discussion of exemplary solutions after the remaining time within the plenary group takes place.

4.2.2 Critical-Thinking Pattern (CTP)

The *group goal* of the CTP is that learners simultaneously and collaboratively correct and improve an existing solution from a subtask within two hours. The collaboration helps them to achieve HLL effects in the discipline of critical thinking as well as teamwork and communication abilities. Typically, abstract solutions of HLL knowledge look professional, complex and thus, seem to be correct. Hence, the *group product* is an improved solution comprising text and visualizations for each subtask. This leads to subtasks that constitute sample solutions that challenge the learners in a way that HLL on the upper levels of Bloom's taxonomy (analyze, evaluate, create) will be addressed. The CTP comprises three distinct steps, each using a thinkLet to structure *group activities* to improve an existing solution. While the learners work within a subgroup in step 1, they collaborate in step 2 and rate their results individually in step 3. In step 1, each learner receives access to the existing solutions of all subtasks. On their own, each learner analyzes all provided solutions, marks mistakes, and makes notes for improvements. In step 2, learners are assigned to breakout groups. Each receives a subtask solution with a list of marked mistakes and improvements. Within breakout groups learners evaluate, interpret, and explain the solutions. They clarify improvement suggestions and write down a revised solution in the form of text and visualizations. A member of each breakout group presents the revised solution to the subgroup. In step 3, learners evaluate on their own whether the solutions of the subtasks are correct and whether they are satisfied with it. The *tool* support in steps 1 to 3 provides similar collaborative working spaces as the PSP with shared writing pages. Here, learners can mark mistakes (step 1), create a revised solution (step 2), and rate the revised solutions (step 3). After each step the tools generate a report of the current deliverables of the step (e.g. list of mistakes). The *group behavior* is restricted toward a focused collaboration like in the PSP. Table 4 illustrates the CTP and serves as moderation plan.

Table 4. Critical-Thinking Pattern (CTP): Overview and Moderation Plan

Learning objective	Analyze, evaluate, and create.	**Task specification** Content, context, connection, appeal to learners. Subtasks are exemplary solutions with mistakes.
Individual goal	Training HLL in the discipline of critical thinking, and teamwork and, communication abilities.	
Group product	Correct and abstract solution in the form of text and/or visualization.	
Group changes	1: subgroup to breakout group	
Tool support	GSS – separate groups, text editing, visualizations, voting: e.g., ThinkTank, GoogleDocs; flip charts, post-its, cards	

	Group activity & general description	Group product & quality indicator	Group procedures (thinkLet, PoC) [8]
Step 1: 30 min	Subgroup: Presentation of the solution of all subtasks and identification of mistakes and inconsistencies.	Product: Per subgroup, a list with identified mistakes and suggestions for improvements for each subtask. Quality: Each mistake comprises a constructive suggestion for improvement.	BucketWalk (evaluate): For each subgroup a bundle of shared writing pages, each with a subtask solution. a) Learners read and walk through the subtask solutions. b) Learners mark aspects that are: false/not relevant, redundant/inconsistent, poorly formulated. c) Learners write down suggestions for improvements (e.g., better formulation, correct knowledge concepts).
Step 2: 80 min	Breakout group & subgroup: Correction of mistakes and inconsistencies. Finalization of correct solution and its presentation.	Product: Per breakout group, revised mistakes and inconsistencies of the solution. Quality: Revised solution comprises all relevant knowledge concepts in form of text and meaningful visualizations.	BucketBriefing (clarify) : For each breakout group a shared writing page for a subtask. a) Learners discuss and clarify marked mistakes and improvement suggestions. b) Learners write down a revised solution for their subtask. c) A member of each breakout group presents revised solution in front of the subgroup (max. 5 min).
Step 3: 10 min	Breakout group: Assessment of the final solution by learners.	Product: Per breakout group, a rated solution of every subtask. Quality: Positive values for correctness and satisfaction with the revised solutions.	MultiCriteria (evaluate): For each breakout group a shared voting page, each for a subtask. a) Post a list of evaluation criteria (level of correctness, level of sophistication, satisfaction with revised solution) for each subtask solution. b) Learners rate each subtask solution on a scale from 1[very bad] to 7 [very good].

5 Evaluation of Collaborative-Learning-Pattern Approach

5.1 Research Method

Data Collection and Measures: We started in 2014 and iteratively designed and evaluated the CLPA using a mixed method approach in line with collaboration engineering to evaluate our design goals [31] (Table 5). We raised explorative findings with real stakeholders and based the evaluation on qualitative and quantitative data [32] that comprised simulations [requirement-based evaluation and identification of stumbling blocks], walk-throughs [interview for stumbling blocks in the process design] by lecturers, and pilot-tests [survey, pre/post knowledge test] by learners (see Table 5). Based on established scales, measures were adapted from Petter et al. 2010 (*plausible; effective; feasible; predictive; reliable*) [9] to build the category system for a content analysis to analyze the qualitative data; and from Briggs et al. 2013 (5-item scales – *satisfaction with process [SP]; satisfaction with outcome [SO]; tool difficulty [TOOLDIF]; process difficulty [PROCDIF]*) [8]. Moreover, we used a pre/post knowledge test, each comprising five single-choice questions to investigate findings for knowledge increases among the learners for each treatment.

Table 5. Mixed-Method Approach to Evaluate the CLPA

	Iterative evaluations	1st	2nd	3rd
Qual. data	Simulation [requirement-based evaluation] (by designer)	N = 1	N = 1	N = 1
	Walk-through [interview] (by lecturers)	-	N = 4	N = 2
Quant. data	Pilot-test [survey, pre/post knowledge test (by learners)	-	-	N = 36

Context and background of the study / participants: All independent lecturers participating in the study teach information systems courses at master levels. The pilot tests were conducted in the same master's course on the topic of "Collaboration Procedures" and thus, with the same tasks. The participants were students from German and Swiss universities. In each semester the course was usually attended by 10 to 20 students. Among all pilot-tests a total of N = 36 students [17 males, 19 females], aged 22 to 34 years [mw = 26 years], participated in the CLPA. The CLPA with IT-supported tools (ThinkTank) and paper-based tools (flip chart, cards) was conducted by us as designers and by two lecturers, leading to four subgroups, each representing a treatment (Figure 2). *Procedures:* Before the evaluation in the field, the quality of the CLPA was assessed using a requirement-based evaluation by us as designers to investigate whether the design of CLPA meets the generalizable requirements. During the walk-through, the design of the CLPA was presented to lecturers and they were asked to identify inconsistencies. Participating in the pilot-test was voluntary and served as preparation for the final exam of the course. The two patterns were bundled, which created a 5-hour learning experience. Learners of a subgroup received a problem situation with four subtasks that required them to describe a blueprint of effective collaboration in the form of a storyline with scenes; each scene had to be described in an abstract and sophisticated way using text and visualizations to demonstrate knowledge concepts. The four subtasks constituted several sequences of scenes. First, learners completed a pre knowledge test, then passed the PSP and CLT, and finally completed a post knowledge test and a survey.

		CLPA experience	
		paper-based tools	IT-supported tools
Moder-ation	Designer	Treatment 1: [group A]	Treatment 2: [group B]
	Lecturer	Treatment 3: [group C]	Treatment 4: [group D]

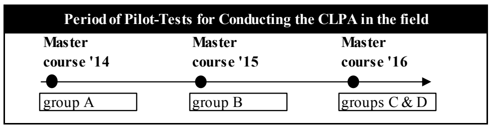


Figure 2. Treatments in the Pilot-Tests

5.2 Results

Simulation by designers and walk-throughs with lecturers. We used a content analysis based on Kohlbacher 2006 [32] and grounded a category system on measures for pattern evaluation based on Petter et al. 2010 [9] (Table 5). *1st evaluation:* Plausibility, effectivity, and feasibility were examined by a simulation. There were no inconsistencies. To judge whether the CLPA is predictive or reliable was not possible. *2nd evaluation:* Walk-throughs with lecturers resulted in statements such as *"when do the learners work in groups and when do they work alone"*. We refined the comprehensiveness of instructions for the lecturer and the subtask wording for learners to improve effectivity. We also refined the grouping structure to improve instructions and rewrote its wording. With regard to the question whether the CLPA is predictive and reliable, the lecturers felt comfortable and were sure that *"the activities will work and the learners will be motivated"*. *3rd evaluation*: A lecturer stated his *"[…] feeling of being a coach"*. The discussion with each lecturer was, inter alia, about whether the process design of CLPA was effective and whether it was reliable. With the help of statements like *"[…] whether the time of that activity is realistic,* *depends upon the number of subtasks […]"* or *"how does that subsolution serve as relevant input for the next subtask; what are input-output relations between the subtasks?"* we improved the time and the sequence of activities, and thus the granularity of activities. We bundled activities to blocks and adapted validated thinkLets from CE. Moreover, the requirement-based evaluation was in line with Hevner 2004 and 2007 [11, 33], and indicated that the process designs of CLPA met the generalizable requirements and thus, coped with the demands of the environment and the body of CL literature. *Pilot-Tests with learners* helped examine whether CLPA met the design goals. We derived three hypotheses, each with exploratory research questions that guided our data analysis:

$H_{1:}$ The CLPA conducted by the designer results in high learner satisfaction.

- $Q_{1a:}$ Did CLPA with paper-based tools result in high learner satisfaction (T1)?
- $Q_{1b:}$ Did CLPA with IT-supported tools result in high learner satisfaction (T2)?

$H_{2:}$ Lecturers are able to conduct the CLPA as good as the designer of the CLPA, so that learners are equally satisfied regardless of the moderator.

- $Q_{2a:}$ Did conduction of CLPA by different moderators and with the same paper-based tools result in similar learner satisfaction comparing treatment 1 and 3?
- $Q_{2b:}$ Did conduction of CLPA by different moderators and with the same IT-supported tools result in similar learner satisfaction comparing treatment 2 and 4?

$H_{3:}$ The conduction of the CLPA with different tool support leads to comparable scores of perceived satisfaction by the learners.

- $Q_{3a:}$ Did conduction of the CLPA by the designer and with different tool support lead to a difference in learner satisfaction in treatment 1 and 2?
- $Q_{3b:}$ Did conduction of the CLPA by lecturers and with different tool support lead to a difference in learner satisfaction in treatment 3 and 4?

To make sure that groups started with no bias with regard to group size, gender, and age, we ran a Kruskal-Wallis test. The results showed no significant difference. To investigate findings for design goal 1 with regard to knowledge increases we compared the means of pre/post-knowledge tests in all treatments. There was a significant difference in the knowledge test performance in each treatment. Learners performed better in the post knowledge test (mean = 3.6) than in the pre knowledge test (mean = 3.0) (Table 6). To verify whether the construct scores have a better mean than a test score (neutral average score on 7-point Likert scale) we run a 1-sided t-test [34] to examine H1. The analysis of Q1a and Q1b showed that all constructs differed significantly, except in terms of TOOLDF for Q1b. Means were better than the average test score and thus on average and upper levels of the 7-point Likert scale (Table 7). To analyse H_2 and H_3, we run a Mann-Whitney test. The results indicate that learners rated the satisfaction in all treatments on upper levels. To investigate H_2, we analysed whether the CLPA can be conducted by different moderators (designer vs. lecturer). Q_{2a} focused on the paper-based tool conduction of the CLPA by different moderators. We compared the means from treatment 1 and 3. There is no significant difference in the means of SP, SO, and PROCDIF. However, for TOOLDIF ($p<0.000$)

learners in treatment 1 (mean = 6.138) scored significantly higher than learners in treatment 3 (mean = 4.100). Q_{2b} focused on the IT-supported CLPA conduction by different moderators. Thus, we compared the means from treatment 2 and 4. There is no significant difference in the means of SP, SO, TOOLDIF, and PROCDIF. To investigate H_3, we analysed whether the CLPA can be conducted with different tool support (paper-based tools vs. IT-supported). Q_{3a} focused on the CLPA conduction by a designer with different tool support. A comparison of the means from treatment 1 and 2 showed no significant difference for SP, SO, and PROCDIF. However, for TOOLDIF there was a significant difference by treatment ($p<0.000$). Learners in the paper-based treatment 1 (mean = 6.138) scored significantly higher than learners in IT-supported treatment 2 (mean = 4.089). Q_{3b} focused on the conduction with different tool support by lecturers. There is no significant difference for SP, TOOLDIF, and PROCDIF when comparing the means from treatment 3 and 4. But learners in the paper-based treatment 3 (mean = 5.640) scored significantly higher than learners in the IT-supported treatment (mean = 6.514).

Table 6. Subgroup Structure: Manipulation Check and Knowledge Increases

	N	gender: male	gender: female	age	pre-test knowledge	post-test knowledge	p-value (2-tailed)
all groups	36	17	19	26	3	3,6	0.000**
group A	8	5	3	28	3,1	3,6	0.033*
group B	10	7	3	26	2,9	3,7	0.003**
group C	11	5	6	25	2,9	3,5	0.011**
group D	7	0	7	25	2,7	3,5	0.045*
p-value (2-tailed)	1.000 ns	0.031 *		0.175 ns	0.321 ns	0.846 ns	-

Note: Kruskal-Wallis test; mean difference significant **p< 0.01, *p < 0.05, ns = not significant; knowledge test (5-item scale)

Table 7. Evaluation Results: Means, Differences in Satisfaction

	Treatment 1 (DP)		Treatment 2 (DI)		Treatment 3 (LP)		Treatment 4 (LI)		Q_{1a} T1	Q_{1b} T2	Q_{2a} T1 vs.T3	Q_{2b} T2 vs. T4	Q_{3a} T1 vs. T2	Q_{3b} T3 vs. T4
	group A		group B		group C		group D		t-value (1-tailed)	t-value (1-tailed)	t-value (2-tailed)	t-value (2-tailed)	t-value (2-tailed)	t-value (2-tailed)
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)						
SP	8	5.988 (0.66)	9	5.822 (0.86)	10	5.940 (0.74)	7	6.029 (0.51)	0.000 **	0.000 **	0.929 ns	0.470 ns	0.606 ns	0.669 ns
SO	8	6.025 (0.68)	9	5.533 (1.24)	10	5.640 (0.76)	7	6.514 (0.50)	0.000 **	0.003 **	0.474 ns	0.055 ns	0.606 ns	0.025 *
TOOLDIF	8	6.138 (0.71)	9	4.089 (0.76)	10	4.100 (0.54)	7	3.714 (0.45)	0.000 **	0.368 ns	0.000 **	0.174 ns	0.000 **	0.133 ns
PROCDIF	8	5.163 (0.91)	9	5.756 (0.59)	10	5.680 (0.61)	7	5.486 (0.28)	0.005 **	0.000 **	0.081 ns	0.210 ns	0.093 ns	0.536 ns

Note: Mann-Whitney test; 7-point Likert scale (1= very less; 7 = very high); mean difference significant **p< 0.01, *p < 0.05, ns = not significant

6 Discussion

In the following, we discuss the results with respect to the two design goals defined at the outset of this paper. *DG 1 – help lecturers to conduct CL activities for HLL in the classroom:* Results from the qualitative content analysis provided insights on how to improve the design of the CLPA. Two lecturers conducted CLPA during several pilot schemes and achieved comparable results (increases in knowledge test performance; satisfaction measures) with the learners compared to the conduction by the designer. The results regarding H_2 with Q_{2a} and Q_{2b} showed no significant difference in the scores; except for TOOLDIF in Q_{2a}. The difference in the TOOLDIF may indicate that use of paper-based tool support should be described in more detail. The results show that lecturers become empowered to conduct CLPA and that the CLPA has the potential to enable knowledge increase among learners. *DG 2 – help learners to proceed through CL activities for HLL:* Among all treatments the satisfaction scores were above an average score of 4 and thus, on average and upper levels of the 7-point Likert scale (H_1). This indicates that learners are able to follow the CL activities in a positive manner. H_3 focused on whether there is a difference in the conduction of CLPA with paper-based tools and IT-supported tools and thus, which way of tool support is easier for learners to follow. To avoid bias by moderator we compared treatments 1 and 2 (both moderated by the designer) to gain insights for Q_{3a}; and treatments 3 and 4 (both moderated by a lecturer) to gain insights for Q_{3b}. Q_{3a} showed a significant difference in the measures of TOOLDIF ($p<0.000$). Learners felt more comfortable with paper-based tool support, since they may have perceived the collaboration as being closer. Another explanation could be that they perceived visualizing or editing contributions as a more flexible way, and thus felt more comfortable with it. However, comparing lecturer moderated treatment 3 and 4 results showed no significant difference in TOOLDIF. Thus, the difference in treatment 1 and 2 may be attributed to the facilitation skills of the designer who moderated the CLPA experience. A similar conclusion can be drawn with regard to Q_{3b} for SO ($p<0.025$). Learners in IT-supported treatment 4 are more satisfied with the outcome than learners in paper-based treatment 3. Thus, the SO with IT-supported tools seems to be more satisfying. However, when comparing treatment 1 and 2, there is no significant difference in means of SO. An explanation for the significant difference of SO in treatment 3 and 4 may be attributed to the facilitation skills of the lecturer.

7 Limitations, Future Research, Contribution, and Conclusion

This study is not without limitations, which provide future research opportunities. First, the evaluation of CLPA was communicated as a HLL experience. For that reason we built exemplary instances that bundled the PSP and the CTP. Consequently, learners followed a HLL experience in which they passed the PSP and then the CTP. It would be valuable for future research to evaluate each pattern on its own. Second, in total we have N = 36 learners that participated in the CLPA (four subgroups). To strengthen our results, it would be valuable for future research to evaluate the CLPA with more groups in a large-scale lecture. Thirdly, the design goals of this study referred to enhancing lecturers to conduct CL activities for HLL and to providing learners guidance to proceed through these activities. The focus was not on evaluating learning success. Hence, future research should investigate knowledge increases among learners in more detail – e.g. group deliverable evaluations by independent lecturers. In particular, follow-up evaluation will need to assess critical thinking and problem-solving skills in more detail. The contributions of the study are positioned along the components of DSR: The purpose and scope of the CLPA is to package sufficient collaboration expertise to conduct CL activities for HLL. To address this set of unsolved problems we provide principles of form and function inherent in generalizable requirements and the CLPA design with its two patterns. This provides guidelines to enable CL activities for HLL in the classroom.

We outlined the CLPA as an approach that helps lecturers to leverage the power of HLL in the disciplines of problem-solving and critical thinking. We based our research on justificatory knowledge from CL and collaboration engineering, and thus, postulate CLPA's potential for enhancing HLL. With three design and evaluation cycles we build expository instantiations of CLPA and evaluated it with real stakeholders by using a mixed methods approach. The results provide insights for CL literature since they show that principles from CE literature can be applied to the field of learning in a way that process restrictions have the potential to support learners in their HLL experience. With the CLPA design we provide insights on how to design CL activities that package sufficient collaboration expertise to empower lecturers to conduct those activities in a predictive way and provide learners guidance to cope with open-ended HLL tasks. Scalability of CLPA is given when several subgroups (with breakout groups) work simultaneously, since CL activities take place there. The CLPA provides prescriptive knowledge and resembles a *'theory of design and action'* [13] of the contribution type *'improvement'* [12].

Acknowledgements

Research presented in this paper was funded by the German BMBF in course of the project StaySmart (www.projekt-staysmart.de), FKZ 01FK14008.

References

1. Krathwohl, D.R.: A Revision of Bloom's Taxonomy: An Ovierview. Theory Into Practice 41, 212-218 (2002)
2. Chiru, C., Ciuchete, S.G., Lefter, G.G., Paduretu, E.: A Cross Country Study on University Graduates Key Competencies. An Employer's Perspective. Procedia - Social and Behavioral Sciences 46, 4258-4262 (2012)
3. Oeste, S., Lehmann, K., Janson, A., Leimeister, J.M.: Flipping the IS Classroom – Theory-Driven Design for Large-Scale Lectures. 35th International Conference on Information Systems, vol. 35, pp. 1-12, Auckland, New Zealand (2014)
4. Ma, J., Baum, S., Pender, M., Bell, D.W.: Trends in College Pricing 2015. The College Board (2015)
5. Jones, M.G., Brader-Araje, L.: The Impact of Constructivism on Education: Language, Discourse, and Meaning. American Communication Journal 5, (2002)
6. Dillenbourg, P.: Over-Scripting CSCL: The Risks of Blending Collaborative Learning with Instructional Design. In: Kirschner, P.A. (ed.) Three Worlds of CSCL. Can We Support CSCL? Open Universiteit Nederland, Heerlen (2002)
7. Kollar, I., Fischer, F., Hesse, F.W.: Collaboration Scripts - A Conceptual Analysis. Educational Psychology Review 18, 159 - 185 (2006)
8. Briggs, R.O., Kolfschoten, G.L., de Vreede, G.-J., Lukosch, S., C., A.C.: Facilitator-in-a-Box: Process Support Applications to Help Practitioners Realize the Potential of Collaboration Technology. JMIS 29, 159-193 (2013)
9. Petter, S., Khazanchi, D., Murphy, J.D.: A Design Schience Based Evaluation Framework for Patterns. The DATA BASE for Advances in Information Systems 41, 9-26 (2010)

10. Briggs, R.O., Kolfschoten, G.L., de Vreede, G.-J., Albrecht, C., Lukosch, S., Dean, D.L.: A Six-Layer Model of Collaboration. In: Jay F. Nunamaker Jr., Nicholas C. Romano Jr., Briggs, R.O. (eds.) Collaboration Systems, pp. 221-228. Advances in Management Information Systems New York (2014)
11. Hevner, A.R.: A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems 19, 87-92 (2007)
12. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly 37, 337-355 (2013)
13. Gregor, S.: The Nature of Theory in Information Systems. MISQ 30, 611-642 (2006)
14. Topping, K.J.: Trends in Peer Learning. Educational Psychology 25, 631-645 (2005)
15. Moll, L.C.: L.S. Vygotsky and Education. Taylor & Francis, New York (2013)
16. Eisenkopf, G.: Peer Effects, Motivation, and Learning. Economics of Education Review 29, 364-374 (2010)
17. Moore, M.G., Kearsley, G.: Distance education: A systems view of online learning. Wadsworth Publishing Company, Belmont, California (2011)
18. Jones, J.M.: Discussion Group Effectiveness is Related to Critical Thinking through Interest and Engagement Psychology Learning and Teaching 13, 12-24 (2014)
19. Wegener, R., Leimeister, J.M.: Peer Creation of E-Learning Materials to Enhance Learning Success and Satisfaction in an Information Systems Course. 20th European Conference on Information Systems, vol. 20, pp. 1-12, Barcelona, Spain (2012)
20. Parece, J., Mulder, R., Baik, C.: Involving Students in Peer Review Case Studies and Practical Strategies for University Teaching. Centre for the Study of Higher Ed. (2009)
21. Harris, A.: Effective Teaching: A Review of the Literature. School Leadership & Management: Formerly School Organisation 18, 169-183 (1998)
22. Hall, T., Stegila, A.: Peer Mediated Instruction and Intervention. NCAC Classroom Practices (2003)
23. Mohammed, S., Dumville, B.C.: Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. Journal of organizational Behavior 22, 89-106 (2001)
24. Bossche, P.V.d., Gijselaers, W., Segers, M., Woltjer, G., Kirschner, P.: Team learning: building shared mental models. Instructional Science 39, 283-301 (2010)
25. Barrows, H.S.: A taxonomy of problem-based learning. Med. educ. 20, 481-486 (1986)
26. Facione, P.A.: Critical thinking. Retrieved June 9, 2004 (1998)
27. Hmelo-Silver, C.E.: Problem-Based Learning: What and How Do Students Learn? Educational Psychology Review 16, 235-266 (2004)
28. Kirschner, P.A., Sweller, J., Clark, R.E.: Why minimal guidance during instruction does not work. Educational Psychologist 41, 75-86 (2006)
29. Hung, W.: The 9-step problem design process for problem-based learning: Application of the 3C3R model. Educational Research Review 4, 118-141 (2009)
30. Gallupe, R.B., Dennis, A.R., Cooper, W.H., Valacich, J.S., Bastianutti, L.M., Nunamaker, J.F.: Electronic brainstorming and group size. AMJ 35, 350-369 (1992)
31. Leimeister, J.M.: Collaboration Engineering. Springer Gabler, Berlin Heidelberg (2014)
32. Kohlbacher, F.: The Use of Qualitative Content Analysis in Case Study Research. Forum: Qualitative Social Research 7, 1-30 (2006)
33. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28, 75-105 (2004)
34. Lehmann, K., Söllner, M., Leimeister, J.M.: Der Wert von IT-gestütztem Peer Assessment. WI. Osnabrück, Germany (2015)

Using Elected Elements in Large-Scale Information Systems Lectures

Christian Karl Grund[1], Maria Tulis[2]

[1]University of Augsburg, Faculty of Business and Economics, Augsburg, Germany
[2]University of Augsburg, Department of Psychology, Augsburg, Germany
{christian.grund@wiwi,maria.tulis@phil}.uni-augsburg.de

Abstract. Information systems (IS) lectures often address audiences that consist of over one hundred students. In this setting, it is arguably difficult to consider the individual interests of each participant. This may result in students not being motivated, decreased learning outcomes as well as an overall low effectiveness of IS lectures. Self-determination theory suggests that perceived autonomy increases intrinsic motivation, which may in turn lead to improved learning outcomes. We therefore propose to foster perceived autonomy among students by introducing elected elements (e.g., practical examples and topics) that students can vote for with an audience response system. To investigate this instructional approach and to provide an instrument for its evaluation, we conducted a preliminary study that shows positive associations between perceived autonomy, intrinsic motivation, as well as acceptance among students. Based on these findings, we derive several avenues for future research regarding the use of elected elements in large-scale IS lectures.

Keywords: Information Systems Lectures, IS Curriculum, Self-Determination Theory, Audience Response Systems.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Grund, C. K.; Tulis, M. (2017): Using Elected Elements in Large-Scale Information Systems Lectures, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 776-787

1 Introduction

Undergraduate enrollment in degree-granting postsecondary institutions has continuously grown over the past decades. For instance, in the year 2013, this enrollment has increased by 46 percent in the United States compared to the year 1990 [1]. This growth results in information systems (IS) lectures that often address audiences consisting of over one hundred students that passively listen to instructors [2]. In this setting, it is arguably difficult to meet the individual interests of each participant for instance in terms of how the knowledge is embedded in practical examples. This may result in students not being motivated, decreased learning outcomes as well as an overall low effectiveness of IS lectures [3, 4]. Self-determination theory suggests that a possible way to foster students' intrinsic motivation is increasing their perceived autonomy [5], i.e. their perception of being able to choose topics and to influence the course of the lectures. However, asking each and every student about how the lectures should unfold is practically impossible in large-scale lectures.

We therefore propose to use pre-fabricated elements (e.g., practical examples, topics, etc.) that students can choose from in every lecture by voting in an audience response system (ARS). The main idea is that every lecture contains both mandatory elements to ensure certain learning outcomes are met as well as elected elements that meet students' interests and provide a feeling of influence on the course. We therefore pose the following overarching question of our research project:

RQ: What are the impacts of providing elected elements in large-scale information systems lectures on students' intrinsic motivation and learning outcomes?

In this paper, we report the results of a preliminary, cross-sectional study conducted in an introductory IS course at a German university, where students were given the choice over several elected elements in each lecture. After the course was finished, different aspects of intrinsic motivation and perceived autonomy were assessed with a brief questionnaire based on the intrinsic motivation inventory [6]. In the present study we were primarily interested in (a) analyzing the psychometric properties of the items and respective scales to provide a reasonable and valid measurement for a subsequent quasi-experimental field study, (b) exploring the acceptance and practicability of the instructional approach by gathering students' qualitative feedback as well as their ratings on an additional "Desirability" scale, and (c) providing first indications regarding its motivational benefits by examining qualitative feedback and performing correlational analyses of students' self-reported perceived autonomy, intrinsic motivation (in terms of interest/enjoyment), and subjective value (in terms of perceived usefulness) of elected elements. Based on the theoretical assumptions of self-determination theory, we expected positive associations between perceived autonomy and intrinsic motivation on the one hand, and intrinsic motivation and perceived value of elected elements. Due to privacy concerns, we were not able to collect performance data (i.e., learning outcomes) in the present study, which will be included in the subsequent study by using anonymous ID-codes.

The remainder of this paper is organized as follows: First, we provide the theoretical background for this study as well as related work. We then report the setup, method, and results of the preliminary study. Afterwards, these results are discussed and avenues for future research are shown in the concluding section.

2 Theoretical Background and Related Work

Self-determination theory stems from motivational psychology and provides several explanations for human motivation [5]. One of its central assumptions is that intrinsic motivation (i.e., the highest level of self-determination; when individuals engage in behavior for the pleasure and satisfaction that they inherently experience with participation [5]) requires the satisfaction of three basic psychological needs: Perceived competence, relatedness, and autonomy [7]. While perceived competence in lectures is already addressed by approaches that test knowledge and understanding of students [8] and relatedness might be covered with peer-reviewing activities [9], perceived autonomy (i.e., being able to influence the course of the lectures) is still rarely addressed by existing studies. While achieving intrinsic motivation among students is one goal of higher education, increased motivation should also lead to better learning outcomes. Indeed, several studies have provided evidence for a link between intrinsic motivation and learning outcomes, such as improved grades [10] or high academic performance through increased study effort and deep learning [11].

By using ARS, which are sometimes also called Audience Response Technology, Personal Response Systems, Electronic Voting Systems or simply “clickers” [12], students may participate in votes with electronic devices. Depending on the infrastructure of the institution (e.g., wireless LAN), this approach may involve many participants [2], which makes it applicable in large-scale lectures (100+ students) as well as in smaller lectures. In addition, studies show that technology-savvy students appreciate ARS, which indicates its usefulness in IS lectures [8]. Several different electronic voting mechanisms have been proposed and used thus far. One popular approach is to distribute designated voting devices to students which they sometimes also have to purchase [8]. However, since the advent of smart phones and tablets, ARS that allow students to use their own devices promise to lower expenses on infrastructure [13]. For this reason, we used such an ARS in the present study. Previous studies that investigated the use of ARS to alleviate the consequences of passive listening in large-scale lectures reported increased engagement [14–17], increased overall satisfaction of students [18–20] as well as increased learning outcomes [14, 21, 22]. However, most of these studies only use ARS to test knowledge of students [16, 17, 22] or to ask for their opinions regarding the content [8]. Only one approach we found in literature might facilitate perceived autonomy by utilizing so-called “clicker cases”, where ARS were used by students to choose several actions in a case study [14]. Although this approach shows how ARS can lead to improved participation, the authors did not examine whether these choices actually had an impact on perceived autonomy of students. Since this theoretical lens may increase our understanding of ways to foster students’ perceived autonomy and ultimately intrinsic motivation in large-scale IS lectures, we focused on perceived autonomy and its associations with other motivational constructs that are described in the following. It is important to note, however, that in this preliminary study we solely address these motivational aspects and the practicability of this specific instructional approach. On the basis of the present findings (improved measurements and instructional approach), the effects of elected elements will be examined in a subsequent quasi-experimental field study with "using elected elements" as independent variable and students' motivation and performance as dependent variables.

3 Preliminary Study

3.1 Implementing Elected Elements in IS Lectures

To investigate the associations between providing elected elements in large-scale IS lectures and students' perceived autonomy as well as intrinsic motivation, we implemented such elements in an introductory IS course at a German university. The course consisted of 12 lectures that were given weekly over a period of 6 months. At the end of each lecture, students were able to vote which element they wanted to be addressed in the following lecture out of 2-4 options. To foster student participation, we used an ARS that allowed students to use their mobile devices (e.g., smart phones, tablets) for voting anonymously [13]. Figure 1 provides a visualization of lectures incorporating elected elements.

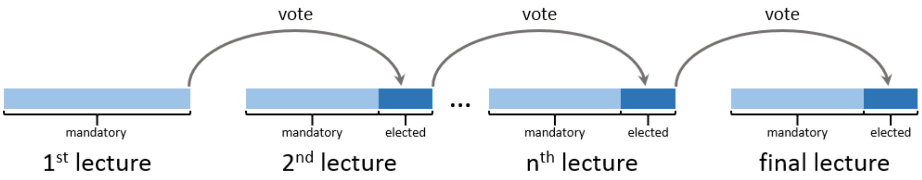


Figure 1. Lectures incorporating elected elements

The elected elements ranged from choosing between different practical examples to choosing between different software demonstrations. For instance, one week before the lecture about business process modeling took place, students were able to choose between activity diagrams and business process model and notation (BPMN) as additional modeling notations. Although these notations are quite similar regarding how they depict business processes, students may get a feeling to be able to choose between a more universal notation (activity diagram) and a notation specifically designed for business processes (BPMN). This way, certain learning outcomes may be enforced while still providing a sense of influence. After voting, students were able to see the distribution of votes between the elected elements. They hence received immediate feedback whether their vote belonged to the majority or not. Due to the fact that every student could participate in many polls, it was very unlikely that they always ended up in the minority, which would arguably reduce their perceived autonomy.

3.2 Method

At the end of the course, after all lectures were finished, we conducted a paper-based questionnaire (see Appendix) in class. Hence, we followed a cross-sectional design in this preliminary study. Data collection was conducted by the first author. Out of the 64 questionnaires we received, 58 have been valid (i.e., 6 were discarded because of obvious dishonesty like wrong fields of study or because they could not participate in the polls due to technical errors with their mobile devices). Since the average number of participants in the votes is 57, the dropout rate appears to be low. Participants consisted of 44 males and 14 females, enrolled in two different fields of study (45 participants studied business and information systems engineering and 13 studied computer science). The age was 20.2 years on average (*SD*=2.7).

The questionnaire adapted items from the intrinsic motivation inventory (IMI) that has been used in many studies to measure perceived autonomy as well as intrinsic motivation of participants [6, 23, 24]. For our preliminary study, we chose three items from each of the following subscales: "Interest/Enjoyment" (e.g., "I would describe the elected elements as very interesting") was used to assess intrinsic motivation, "Perceived Choice" (e.g., "I voted for elected elements because I wanted to") was selected to measure perceived autonomy, and "Value/Usefulness" (e.g., "I think the possibility to vote for elected elements is important") was used to gather an overall rating of subjective value of providing elected elements. Judging an activity to have personal value and importance can be seen as (antecedent) part of intrinsic motivation [5], and therefore it should be positively related to interest and enjoyment. All items were modified to relate to the context and translated into German. However, since these adapted items did not fully cover our research question, we added two self-developed subscales, each comprising three items. "Perceived Influence" (e.g., "By voting for elected elements I felt that I could influence the lectures") addressed an additional aspect of perceived autonomy, since the "Perceived Choice" subscale exclusively asked whether students believed that they participated voluntarily in the polls. However, we also wanted to know whether they believed that their votes had an impact on the lectures. Finally, "Desirability" (e.g., "I wish I had the possibility to vote for elected elements in other courses, too") was added as another way of asking for an overall rating of providing elected elements, since the "Value/Usefulness" subscale only asked whether elected elements are important to students. While this is a possible approach to determine an overall rating, we also wanted to know whether students thought that providing elected elements makes sense and whether they wish having these elements in other courses, too. Every subscale except for "Value/Usefulness" contained one reversely coded item that was used to identify fraudulent questionnaires (i.e., there should be no contradictions). Each item in the questionnaire was assessed using a 5-point scale, ranging from 1 = *not at all true* to 5 = *very true*, and they were randomized across all subscales. In addition to the quantitative items, students were provided with space for leaving any comments or suggestions on the possibility of voting for elected elements. All subscales as well as their respective internal consistencies (Cronbach's α) are presented in Table 1. Reliabilities were satisfactory for all subscales, except

“Perceived Choice”. Thus, the subscale “Perceived Influence” provided a more consistent measurement of perceived autonomy, and “Perceived Choice” was omitted.

Table 1. Subscales and Cronbach’s α

Subscale	Number of Items	Cronbach’s α
Interest/Enjoyment	3	0.79
Perceived Choice	3	0.43
Perceived Influence	3	0.82
Value/Usefulness	3	0.83
Desirability	3	0.83

3.3 Results

Confirmatory factor analysis, used to verify the latent factor structure (i.e., subscales) of the measurement instrument, revealed an acceptable fit for the remaining four subscales (χ^2=68.19, *df*=48, *p*=0.03, CFI=0.95, TLI=0.92, RMSEA=0.08). Standardized item loadings were in the range of λ = 0.67 - 0.85, thus satisfactory. The usual and recommended cut-off scores for RMSEA are ≤ 0.05 for a good fit and ≤ 0.08 for an acceptable fit. CFI and TLI should be ≤ 0.95 for a good fit, and ≤ 0.90 for an acceptable fit [25]. Hence, all further analyses were based on the four subscales “Interest/Enjoyment”, “Perceived Influence”, “Value/Usefulness”, and “Desirability”. The descriptive statistics are shown in Table 2. The item numbers indicate the sequence of questions. Bivariate intercorrelations (manifest) also indicate discriminant validity of the different aspects of motivation (see Table 3).

Table 2. Descriptive statistics of the results (N=58)

Subscale	Item No.	Mean	SD	Power	Loading
Interest/Enjoyment	3	3.76	0.84	0.70	0.78
	12	4.03	0.83	0.63	0.79
	15	3.59	0.74	0.57	0.67
	Total	**3.79**	**0.67**		
Perceived Influence	4	3.60	0.95	0.64	0.70
	7	3.45	0.95	0.69	0.84
	8	3.86	0.95	0.68	0.76
	Total	**3.64**	**0.81**		
Value/Usefulness	9	3.53	0.99	0.60	0.69
	14	3.36	0.96	0.73	0.83
	10	3.40	1.08	0.73	0.85
	Total	**3.43**	**0.87**		
Desirability	5	4.12	0.97	0.72	0.79
	16	3.98	0.88	0.60	0.77
	11	4.40	0.83	0.75	0.81
	Total	**4.17**	**0.77**		

A closer look at the distributions of these subscales is provided in Figure 2.

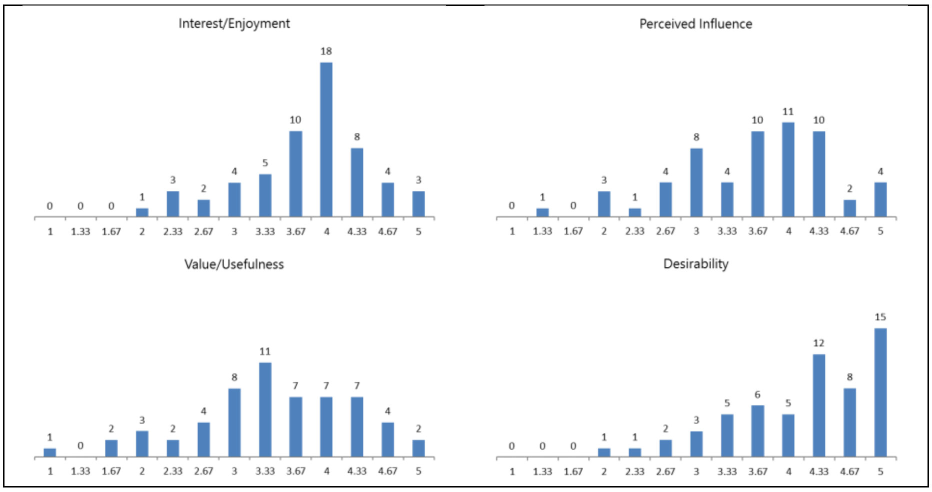


Figure 2. Distribution of each subscale

As Figure 2 shows, particularly the distribution of “Desirability” was more concentrated towards the higher end of the scale (negative skew) whereas the distribution of “Perceived Influence” was rather scattered. Students generally reported rather high levels of interest and enjoyment after using elected elements and they reported a strong desire to have such elements in other courses. Students’ ratings regarding the value and importance of these elements as well as the amount of perceived autonomy in terms of perceived influence they had on the course differed.

Students’ qualitative feedback supports these assumptions by comprising both positive as well as negative comments on the implemented instructional approach. Many students appreciated being able to vote for elected elements in each lecture. Some of the comments also directly state that the interest in the lecture increased by choosing elected elements (45% of all comments):

“By being able to vote for elected elements, one is able to influence the content of the lectures -> increased interest.”

“The interest in the course increases when elected topics are covered.”

“I really liked deciding for the topics that I was most interested in.”

However, there was also criticism regarding how the elected elements were covered in the lectures. The main concern was that they have been too short compared to the mandatory parts of the lectures (36% of all comments):

"The idea of providing elected elements is very good. However, they often have been addressed shortly at the end of the lectures. For example, we were shown how a system from SAP looks, however, I seldom understood how it worked."

"I like being able to vote for elected elements as well as the use of them – However, they have been covered too short in the lectures. When using elected elements, you should take enough time for them."

"Despite the elected elements often being very interesting, they have been covered way too short in the lectures, which made the choices feel pretty pointless."

Additional criticism addressed both the amount of information that was provided before voting for elected elements as well as the unclear relevance of these elements in the examination (18% of all comments):

"More info about the elected elements would have sometimes been useful for better forming an opinion."

"The relevance of the elected elements for the examination has sometimes been unclear (they don't appear in the script and there is no handout)."

In summary, the qualitative feedback emphasized that providing elected elements in IS lectures may lead to increased interest and motivation. It also shows, however, that these elected elements should have more room inside the lecture. Otherwise, they could be perceived as pointless which may reduce perceived autonomy. Finally, students have to be supplied with enough information about each alternative to be able to make a well-informed decision.

Regarding the associations between perceived autonomy and intrinsic motivation due to elected elements in IS lectures, we found a significant positive correlation between students' interest/enjoyment and perceived influence (see Table 3). In addition, we found a positive correlation between interest/enjoyment and perceived value in terms of the rated usefulness of elected elements and the desire to use ARS in other lectures as well. Finally, the latter was positively correlated with perceived influence.

Table 3. Bivariate correlations (**p < 0.01)

	Interest/ Enjoyment	Perceived Influence	Value/ Usefulness	Desirability
Interest/Enjoyment	1			
Perceived Influence	0.55**	1		
Value/Usefulness	0.58**	0.49**	1	
Desirability	0.78**	0.60**	0.62**	1

These positive correlations support our initial expectations concerning the associations between perceived autonomy and intrinsic motivation due to elected elements. We hence propose that fostering perceived autonomy by using elected elements in large-scale IS lectures may have the potential to increase students' motivation in terms of subjective value, interest and enjoyment. These correlational findings provide a basis for future research, more specifically for the intended quasi-experimental study with a comparable student population, to examine causal effects of this instructional approach on students' motivation and achievement.

4 Discussion

The results of this preliminary study indicate that providing elected elements in IS lectures might lead to perceived autonomy and increased intrinsic motivation among students. Our findings, based on qualitative and quantitative data, provide a first step towards understanding the effects of using elected elements in large-scale IS lectures. In addition, these elements are perceived well by the participants. The short-scale measures used in this study proved to be reliable to assess "Interest/Enjoyment", "Value/Usefulness", "Perceived Influence", and "Desirability". Since most students enrolled in IS programs are equipped with mobile devices, they provide a good opportunity to let students vote for their favorite content. Once these elected elements are created by the instructor, they may be used several times and even in several different courses. Because many ARS have been improved over the years, conducting these polls is uncomplicated and arguably fewer effort than for example setting up blended learning scenarios with extensive online content. The present study extends prior research by adding a self-determination theory perspective to explain increased motivation when using ARS by increased perceived autonomy during the lectures.

There are, however, some limitations to this study. First, due to the selected ARS, individual choices of students have not been tracked. We were hence unable to investigate motivational differences between students who often voted like the majority compared to those who did not. This might have been one reason for students' differences in perceived influence. Some students also reported technical problems either with their devices or with the network inside the lecture room. To ensure scalability, an ARS that is able to handle many connections at the same time should be used. According to the comments of students, elected elements should have more room inside each lecture. Indeed, these elements sometimes just comprised 10 minutes inside a 90-minute lecture. We will hence prolong them in future investigations. Another limitation of the study is the lack of a control group. We therefore cannot compare the achieved level of intrinsic motivation from using elected elements with a group that did not use these elements. In addition, performance data of students could not be mapped to the questionnaires due to privacy concerns. Hence, we could not investigate whether those students who reported higher levels of perceived autonomy and intrinsic motivation actually performed better than their peers with lower levels, respectively. In a next step, we will include these aspects in the subsequent study design following a quasi-experimental design.

5 Conclusion

In regard to the findings above, instructors from the IS domain may consider incorporating elected elements into their lectures. When doing so, these elements should noticeably influence the contents of the lectures and students must be supplied with sufficient information about every alternative before voting. Due to the limitations mentioned earlier, this study is only a first step towards understanding the use of elected elements in IS lectures. In the subsequent study, we will track individual choices of each participant to see whether students who have often voted like the majority are more motivated than others. Additionally, performance data of each student will be tracked to investigate effects on student learning. This may include results from examinations as well as other performance indicators, such as regularly performed quizzes. Since offering lectures is often a necessity due to increasing enrollment, our preliminary results highlight one feasible opportunity to improve this experience for both students as well as instructors.

References

1. Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., Wilkinson-Flicker, S., Barmer, A., Velez, E.D.V.: The Condition of Education 2015. U.S. Department of Education, National Center for Education Statistics, Washington, DC (2015)
2. Lehmann, K., Söllner, M.: Theory-driven design of a mobile-learning application to support different interaction types in large-scale lectures. In: Proceedings of the 22nd European Conference on Information Systems (ECIS), pp. 1–12 (2014)
3. Beichner, R.J., Saul, J.M., Allain, R.J., Deardorff, D.L., Abbott, D.S.: Introduction to SCALE-UP: Student-Centered Activities for Large Enrollment University Physics. In: Proceedings of the 2000 Annual Meeting of the American Society for Engineering Education (2000)
4. Eison, J.: Using Active Learning Instructional Strategies to Create Excitement and Enhance Learning. University of South Florida (2010)
5. Deci, E.L., Ryan, R.M.: Intrinsic Motivation and Self-Determination in Human Behavior. Plenum Press, New York (1985)
6. Ryan, R.M.: Control and Information in the Intrapersonal Sphere: An Extension of Cognitive Evaluation Theory. Journal of Personality and Social Psychology 43, 450–461 (1982)
7. Deci, E.L., Ryan, R.M.: The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. Psychological Inquiry 11, 227–268 (2000)
8. MacGeorge, E.L., Homan, S.R., Dunning, J.B., Elmore, D., Bodie, G.D., Evans, E., Khichadia, S., Lichti, S.M., Feng, B., Geddes, B.: Student evaluation of audience response technology in large lecture classes. Educational technology research and development 56, 125–145 (2008)

9. Schlagwein, D.: Students as Reviewers and Lecturers as Editors: The Peer Review with Scaffolded Assignments Model. In: Leidner, D., Ross, J. (eds.) Proceedings of the 36th International Conference on Information Systems (ICIS), pp. 1–9 (2015)
10. Black, A.E., Deci, E.L.: The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. Science Education 84, 740–756 (2000)
11. Kusurkar, R.A., Cate, T.J. ten, Vos, C.M.P., Westers, P., Croiset, G.: How motivation affects academic performance: a structural equation modelling analysis. Advances in Health Sciences Education 18, 57–69 (2013)
12. Moss, K., Crowley, M.: Effective learning in science: The use of personal response systems with a wide range of audiences. Computers & Education 56, 36–43 (2011)
13. Reinhardt, W., Sievers, M., Magenheim, J., Kundisch, D., Herrmann, P., Beutner, M., Zoyke, A.: PINGO: Peer Instruction for Very Large Groups. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D., Hernández-Leo, D. (eds.) 21st Century Learning for 21st Century Skills, pp. 507–512. Springer, Berlin, Heidelberg (2012)
14. Lundeberg, M.A., Kang, H., Wolter, B., delMas, R., Armstrong, N., Borsari, B., Boury, N., Brickman, P., Hannam, K., Heinz, C., et al.: Context matters: increasing understanding with interactive Clicker Case studies. Educational technology research and development 59, 645–671 (2011)
15. Rice, R.E., Bunz, U.: Evaluating a Wireless Course Feedback System: The Role of Demographics, Expertise, Fluency, Competency, and Usage. Studies in Media & Information Literacy Education 6, 1–23 (2006)
16. Fitch, J.L.: Student Feedback in the College Classroom: A Technology Solution. Educational technology research and development 52, 71–77 (2004)
17. Guse, D.M., Zobitz, P.M.: Validation of the audience response system. British Journal of Educational Technology 42, 985–991 (2011)
18. Blackman, M.S., Dooley, P., Kuchinski, B., Chapman, D.: It Worked a Different Way. College Teaching 50, 27–28 (2002)
19. Nicol, D.J., Boyle, J.T.: Peer Instruction versus Class-wide Discussion in Large Classes: A comparison of two interaction methods in the wired classroom. Studies in Higher Education 28, 457–473 (2003)
20. Stuart, S.A.J., Brown, M.I., Draper, S.W.: Using an electronic voting system in logic lectures: one practitioner's application. Journal of Computer Assisted Learning 20, 95–102 (2004)
21. Stratling, R.: The complementary use of audience response systems and online tests to implement repeat testing: A case study. British Journal of Educational Technology (2015)
22. Castillo-Manzano, J.I., Castro-Nuño, M., Sanz Díaz, M.T., Yñiguez, R.: Does pressing a button make it easier to pass an exam? Evaluating the effectiveness of interactive technologies in higher education. British Journal of Educational Technology Online First (2015)

23. McAuley, E., Duncan, T., Tammen, V.V.: Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis. Research Quarterly for Exercise and Sport 60, 48–58 (1989)
24. Plant, R.W., Ryan, R.M.: Intrinsic motivation and the effects of self-consciousness, self-awareness, and ego-involvement: An investigation of internally controlling styles. Journal of Personality 53, 435–449 (1985)
25. Hu, L., Bentler, P.M.: Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 6, 1–55 (1999)

Appendix

Questionnaire. Reversely coded items are marked with (R)

Subscale	No.	Item
Interest/Enjoyment	3	The elected elements have been fun.
Interest/Enjoyment	12	I thought the elected elements have been boring. (R)
Interest/Enjoyment	15	I would describe the elected elements as very interesting.
Perceived Choice	13	I voted for elected elements because I wanted to.
Perceived Choice	6	I felt like I had to vote for elected elements. (R)
Perceived Choice	2	I believe I could choose whether to vote for elected elements or not.
Value/Usefulness	9	I think the possibility to vote for elected elements is important.
Value/Usefulness	14	I believe the possibility to vote for elected elements could be beneficial to me.
Value/Usefulness	10	I believe the possibility to vote for elected elements could be of some value to me.
Perceived Influence	4	By voting for elected elements I felt that I could influence the lectures.
Perceived Influence	7	By voting for elected elements I had the impression of being able to codetermine the contents that have been taught.
Perceived Influence	8	I believe that by voting for elected elements I was unable to influence the lectures. (R)
Desirability	5	I think that the possibility to vote for elected elements makes sense.
Desirability	16	I wish I had the possibility to vote for elected elements in other courses, too.
Desirability	11	The possibility to vote for elected elements should be dropped from the course. (R)
Comment (open ended question)	-	I want to note the following regarding the possibility to vote for elected elements at the end of each lecture.

The Twofold Value of IT-Based Peer Assessment in Management Information Systems Education

Roman Rietsche[1], Katja Lehmann[2], Philipp Haas[1] and Matthias Söllner[1,2]

[1] University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland
{roman.rietsche,philipp.haas,matthias.soellner}@unisg.ch
[2] University of Kassel, Research Center for IS Design (ITeG), Kassel, Germany
{katja.lehmann,soellner}@uni-kassel.de

Abstract. Feedback is one of the most influential factor when it comes to learning success of students. Especially large-scale classes at universities often lack feedback caused by the scarcity of resources. Even though a lack of feedback is problematic across all domains, especially when educating future IS executives, who are supposed to lead team members by providing effective feedback. In this study, we use IT-based peer assessment (ITPA) in a large-scale class to support students with feedback during their learning-process. Specifically, our results show that participating in ITPA lead to an increase in knowledge on the content of the class of about 28% on average. Furthermore, students train their ability to provide feedback, measured twice, self-reported and as quality of the feedback received – increased significantly during the class.

Keywords: IT-based peer assessment, feedback, higher-education, large-scale class, technology-mediated learning

1 Introduction

In the last years, the number of students has been constantly rising while the amount of lecturers stayed steady [1]. Thus, the current condition at universities is that many classes consist of hundreds of students and are taught by one lecturer [2]. One consequence of these large-scale classes is the lack of feedback for students [3]. According to Hatti et al. [4], feedback is the third most influential factor for student learning success. Despite the known importance of feedback in the learning process, the final exam is often the only time students receive feedback in those classes, and this feedback is often aggregated into a single performance score providing hardly any insights into strengths and areas that need to be improved.

We argue that especially future managers need to develop their own ability to provide feedback to become an effective leader in business [5]. Thus, the lack of feedback is even more crucial in this domain, since it a) hinders learners in assessing their current state of knowledge and identifying areas for improvement, and b) restrains them from improving their own ability to provide feedback to other people. A possible solution to overcome both barriers and to ensure that students of large-scale classes have the possibility to train their ability to provide feedback, as well as to receive

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Rietsche, R.; Lehmann, K.; Haas, P.; Söllner, M. (2017): The Twofold Value of IT-Based Peer Assessment in Management Information Systems Education, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 788-791

feedback to assess their current state of knowledge and to adapt their learning process, is IT-based peer assessment (ITPA). To assess the value of using ITPA in large-scale classes, we seek to answer research question of:

RQ1: To what extent does participating in an ITPA help to train the learners' ability to provide feedback?

RQ2: To what extent does participating in an ITPA contribute to the learners' knowledge of the learning contents?

According to Hattie and Timperley [4] feedback is conceptualized as information provided by an agent, for example a lecturer who provides corrective information [4]. Hence, feedback is provided a consequence of students' performance. Whereas the outcome of feedback is an information, specifically relating to the task or process of learning that fills a gap between what is understood and what is aimed to be understood [6].

2 Methodology and ITPA process

ITPA was used in a quasi-experiment with a one-group pretest-posttest design in a large-scale university classes on information systems at Master's level. The quasi-experiment consisted of seven ITPAs, which took place on a regular manner over a time period of one semester and in each participating between 73 and 101 students.

Figure 1 shows the process of an ITPA. All steps shown in the process were conducted within a web-based learn management system (LMS). In the first step each student carries out the self-assessment (SA), using a survey including questions of their perception of their ability to provide feedback. In the second step the student creates the assignment and the LMS sends the completed assignments to the three anonymous peer reviewers – other students of the same class– (the reviewers were randomly selected in each ITPA by the LMS).

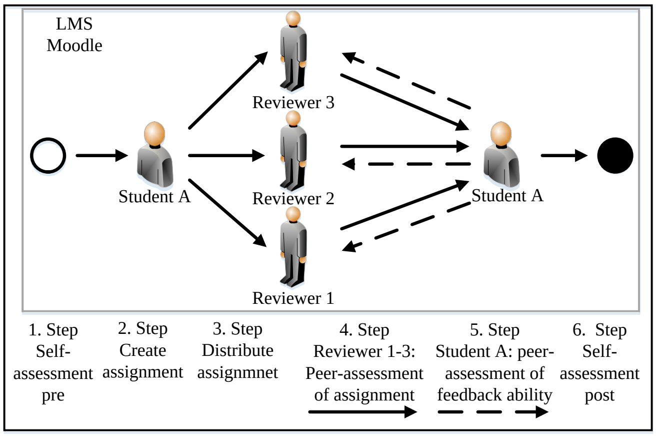


Figure 1. Overview of the ITPA Process

In the fourth step, each reviewer carries out a peer-assessment (PA) of the assignment using a standardized review template. Including questions concerning the strengths and weaknesses of the assignment and how it could be improved. Bevor the students provided their first PA a hands-on instruction of how to provide constructive feedback was carried out. After the PA the reviews are automatically sent back to the respective student by the LMS. In the fifths step, each student revises the initial solution, creates a change history, and rates the quality of the feedback provided by each reviewer based on the same questions as the SA. In the last step, each student carries out another SA. For the post SA the same survey was used as in the pre SA. In case a student did not participate in an ITPA, which leads to a missing value, this was statistically considered (marked as NA in R).

3 Results

The graph in Figure 2 shows the survey results for the question "How do you assess your personal ability to provide feedback to your peers?". The first measurement t0 is the baseline measurement in which the students were asked to self-assess their ability in providing feedback. The data of t0 represents that no ITPA is used at all and t1 that the ITPA is used one time. The students benefited using ITPA one time significantly ($p < .001$). Moreover, using ITPA multiple times, even further enhances the students' ability which shows the increase for SA and PA from t1 to t7 ($p < .05$). Figure 3 shows the survey results of "What is your current state of knowledge concerning the current learning unit?". The results show that students benefit from using ITPA in terms of perceived knowledge growth. The study showed a statistically highly significant increase of in average 28.3 percent from the pre-measurement to the post-measurement of the students' current state of knowledge concerning the current learning unit.

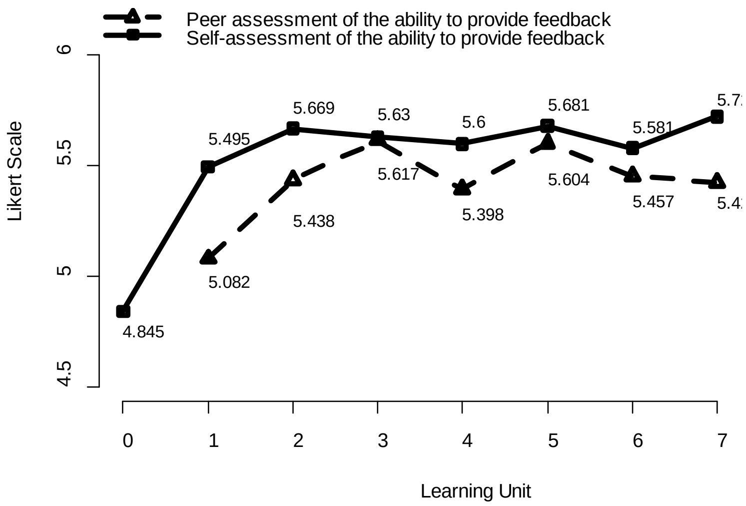


Figure 2. How do you assess your personal ability to provide feedback to your peers?

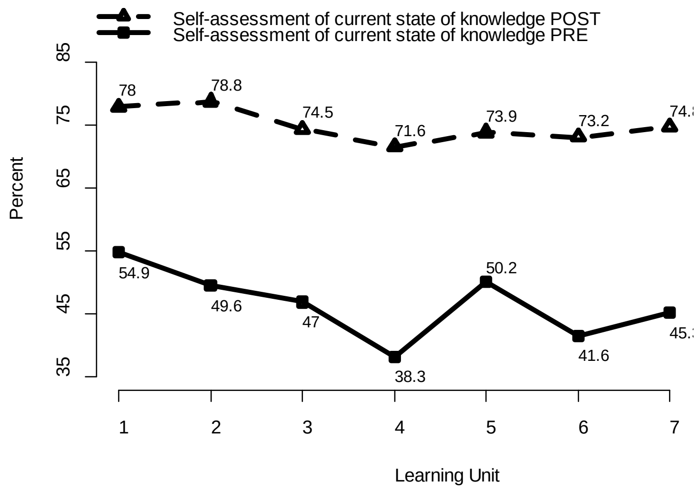


Figure 3. What is your current state of knowledge concerning the current learning unit?

4 Conclusion

This paper describes the usage of ITPA in large-scale classes to train the student's ability to provide feedback. The results show that using ITPA could enable the lecturer to support students in their learning process in a resource-saving way and to provide an environment in which students can anonymously train their ability to provide feedback. Furthermore, the results show that students using ITPA have a perceived increase of knowledge concerning the particular learning unit. With our paper we contribute to the body of literature of feedback in higher-education by providing empirical results from our quasi-experiment.

References

1. Fortes, P.C., Tchantchane, A.: Dealing with Large Classes. A Real Challenge. Procedia - Social and Behavioral Sciences 8, 272–280 (2010)
2. Leidenfrost, B., Strassnig, B., Schabmann, A., Carbon, C.-C.: Verbesserung der Studiensituation für StudienanfängerInnen durch Cascaded Blended Mentoring. Psychologische Rundschau 60, 99–106 (2009)
3. Bligh, D.A.: What's the Use of Lectures? Intellect books (1998)
4. Hattie, J., Timperley, H.: The power of feedback. Review of Educational Research 77, 81–112 (2007)
5. Kaplan, S.N., Klebanov, M.M., Sorensen, M.: Which CEO Characteristics and Abilities Matter? The Journal of Finance 67, 973–1007 (2012)
6. Sadler, D.R.: Formative assessment and the design of instructional systems. Instructional science 18, 119–144 (1989)

Privacy as a Part of the Preference Structure of Users App Buying Decision

Christoph Buck[1], Florian Stadler[2], Kristin Suckau[3], and Torsten Eymann[1]

[1] Chair of Information Systems Management, Bayreuth, Germany
{christoph.buck, torsten.eymann}@uni-bayreuth.de
[2] zeb – 360° Consulting for Financial Services
{florian.stadler}@zeb.de
[3] Chair for Innovation and Dialogue Marketing
{kristin.suckau}@uni-bayreuth.de

Abstract. Information privacy and personal data in information systems are referred to as the 'new oil' of the 21st century. The mass adoption of smart mobile devices, sensor-enabled smart IoT-devices, and mobile applications provide virtually endless possibilities of gathering users' personal information. Previous research suggests that users attribute very little monetary value to their information privacy. The current paper assumes that users are not able to monetize their value of privacy due to its abstract nature and non-transparent context. By defining privacy as a crucial product attribute of mobile applications the authors provide an approach to measure the importance of privacy as part of users' preference structure. The results of the conducted choice-based conjoint Analysis emphasize the high relevance of privacy in users' preference structure when downloading an app and provide an interesting contribution for theory and practice.

Keywords: Information Privacy, Personal Data, Product Attribute, Preference Structure, Mobile Applications.

1 Introduction

With the disruptive innovations of e.g. the iPhone and the iPad software in the form of mobile applications (apps) diffused in the everyday life of users. Apps are integral to the functioning of Smart Mobile Devices (SMD) like smartphones or tablets and are key elements for the interface design and functionality. Apps can be interpreted as the embodiment of ubiquitous computing, i.e. the creation of environments saturated with computing and communication capability, integrated with human users [1]. While ubiquitous computing focuses on hardware components, today's apps are the logical consequence of experiential computing; the "digitally mediated embodied experiences in everyday activities through everyday artifacts with embedded computing capabilities" [2].

At the same time, this development has considerably contributed to the emergence of a new user type of information systems. These new users integrate apps into their

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Buck, C.; Stadler, F.; Suckau, K.; Eymann, T. (2017): Privacy as a Part of the Preference Structure of Users App Buying Decision, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 792-806

everyday lives, which leads to fundamental changes concerning how users interact with computing devices and systems [3].

However, this excessive level of integration does not come without consequences. Many business models are based on the user data collected by SMD, which grants the marketing industry the access to exceptionally valuable information about current and potential customers [4]. Thanks to the mechanics and the real life integration of modern information systems, the value of user data is unique. Thus, users' privacy, increasingly gets at risk.

Given the fact that users' information privacy is a major part of the economic exchange when downloading apps, privacy, and the corresponding settings, have to be determined as an attribute of the value proposition of apps. In order to understand users' concerns and clearly define the necessity of user data protection, it is crucial to determine the value of privacy for users. Caused by the (perceived) abstract nature of personal data and privacy, the current paper states that users are not able to value personal data and privacy in a monetary amount. Consequently, this paper targets users' preference structures when downloading apps. With this in mind, we formulate the following research question:

- Does the protection of privacy, when downloading an app, represent a crucial product attribute for the user?

To answer this research questions, two examinations were conducted both reflecting the importance of privacy as an important attribute when purchasing an app. The HB-based utility value and the estimations of the CBC, show significantly high levels of importance for privacy. In fact, privacy is on first-place in both rankings. The remainder of this article is structured as follows: in the subsequent section, we lay out the groundwork for the definition of privacy as a value and its measurement. Following this, we will describe the methodology conjoint analysis, present our choice based conjoint analysis and its key findings. Finally, we will discuss our findings, address some limitations and conclude with suggestions for further research.

2 The Value of App Privacy

2.1 Information Privacy in the Context of Mobile Applications

Since privacy is addressed in many fields of social sciences and different definitions are used in various areas of everyday life it lacks a holistic definition [4, 5]. First of all, physical and information privacy have to be distinguished. Physical privacy relates to the "access of an individual and/or the individual's surroundings and private space" [4]. Contrary, information privacy only refers to information that is individually identifiable or describes the private informational spheres of an individual. Although information privacy is rooted in the fundamental concept of physical privacy, both are subsumed under the term of "general privacy" [4].

Even though privacy has developed and changed drastically over the last decades, Westin's definition from 1967 still holds true: information privacy is defined as "the claim of an individual to determine what information about himself or herself should be known to others" [6]. Following Westin, 'control' is construed as an instrument of the protection of privacy, that privacy itself is often defined as the control over personal information [5]. Consequently, in this paper information privacy is defined as the ability to control the acquisition and use of one's personal information [7].

As the "pocket knife of communication" [8], SMD possess a vast amount of connected sensors, devices, and functions. SMD in combination with apps are the most common user interface to merge the broad opportunities given by the connected sensors and devices. Throughout these functions, the possibilities of gathering personal data are virtually endless. Future prospects in relation to these applications promise even more opportunities to expand data collection and immediate analysis of data. Regarding data quality, recent developments in mobile technology and an ever-increasing digitization of everyday tasks, lead to an unprecedented precision of continuously updated and integrated personal data, which is generated within mobile ecosystems like iOS and Android [9]. Consequently, apps, as the most common user interface for digitized solutions (e.g., smart services, smart homes, wearables, etc.), layer everyday activities and lives in a digital way; or how Clarke rephrased it: "Cyberspace is invading private space" [10].

In app markets, users are able to control their privacy disclosure during the purchasing process. Thus, users can actively control their disclosure of personal data and the grasping of privacy from third parties [11].

2.2 The Value of App Privacy

Dinev and Hart [12] stated that privacy "is a highly cherished value, few would argue that absolute privacy is unattainable." Privacy as digital personal information and highly personalized data collected via apps has a huge economic value [13]. With the description of personal data as a new asset class, the World Economic Forum [14] is in line with the argumentation of many researchers [4, 15]. Derived from the perspective of personal data and privacy as a commodity [16], many researchers conceive privacy as a tradeable good or asset [15]. According to this view, privacy is no longer an absolute societal value, but has an economic value, which leads to the possibility of a cost-benefit trade-off calculation made by individuals or a society [4].

Nevertheless, the authors of this article argue that privacy cannot be seen as an economic value with (for users) available market prices. First, users' distortion regarding the valuation of their own information privacy is caused by the nature of data collection, aggregation and secondary use of app markets [17]. Following Flender and Müller [18], apps are data-centric services and value is generated on different levels: e.g. between the user and the app provider, the aggregated value of the app as data centric service, and the aggregated data from various apps and underlying ecosystems by third parties [19]. As a result, in app markets it is not possible for users' to reliably evaluate their value of privacy in the moment of releasing personal information. However, major parts of the resulting costs of releasing personal informatio arise by the access, use and transfer of the data on multiple levels. Third parties (e.g., retailers, advertisers, and insurance companies) could, for instance, use that information for issues like price discrimination, advertising or risk surcharges [15]. Accordingly, the value of users' privacy is originated in the release of the information but realized in a sphere, which cannot be controlled by the initial owner (user). In addition, users are often not aware of the possibilities of collection, aggregation and analyzation of digital information [15].

Taking the paradigm of experiential computing into account, the value of privacy increases with the (perceived) invisibility of the connected devices. With the increasing everyday life integration, devices and sensors become more and more invisible but are an increasingly self-evident part of users' daily routine. Because of the establishment in users most intimate privacy sphere, users' awareness regarding their information privacy is affected in a paradox way. In the end, privacy is perceived subjective and individually and the value of different information types and spheres is abstract and intangible. Following these arguments, the presented paper defines privacy as an abstract value. Consequently, users are not able to evaluate the monetary value of their information privacy.

2.3 Related work

When the measurement of the (perceived) value of consumers' information privacy is observed the theory of the privacy calculus has to be considered [20]. Therefore, users are supposed to undertake an anticipatory, rational weighing of risks and benefits when confronted with the decision to disclose personal information [21, 22] or conduct transactions [23]. The privacy calculus model assumes a correct and objectified understanding of the monetary value of privacy and therewith a tangible willingness to pay for privacy of the users [24, 25]. IS privacy research focused on the marketing-based concept of willingness-to-accept (WTA) and willingness-to-pay (WTP) [15]. Although asymmetries and disparities between WTA and WTP have been observed, both concepts are well established in academic research and have been applied to the topic of personal user data multiple times [15, 26]. Besides those disparities and the fact that the ownership of privacy control rights remains difficult to define in the context of apps, WTA and WTP are based on the user's perceived value for privacy and the purchased good or service. As stated above, users are indeed not able to evaluate the monetary value of their privacy, which leads to the impossibility to define the perceived value and thereby the needed maximization or reservation price for WTA or WTP. In the light of the definition of privacy as an abstract value studies which directly elicit users' valuation of privacy in survey settings gain distorted results [27–30] (see for an overview [31]). This is also described by the well-observed phenomena of the privacy paradox [24], which claims that individuals value privacy less than stating in studies and polls. It has been subject of various research in the field of information privacy, but there is no comprehensive explanation why individuals show this paradoxical behavior [24]. Consequently, WTA or WTP do not offer an adequate set of instruments to measure users' privacy concerns or the value they assign to their privacy.

Stemming from that, we recommend the approach of a choice-based conjoint analysis (CBC) to examine users' preference structure when purchasing an app. Some studies measured the preference structure as a proxy for the willingness-to-pay for privacy. Most of these studies are desktop driven and focus on the disclosure on websites and online social networks [31–33] or social app adoption [34]. Despite the increasing studies applying decomposition methods there is no investigation of privacy as a stand-alone product attribute correlated with the provided functionality of the app. Therefore, a CBC is provided which outlines privacy on equal terms to other attributes like price. Hereby we are able to determine if the user only states that he values privacy or if he actually does value it in real life purchase decisions.

Accordingly, a high evaluation of privacy can be assumed when privacy is seen as a crucial product attribute of apps. Therefore, privacy has to be an important product attribute in users' preference structure when buying apps. A well-known and established methodological approach for measuring users' preferences is the Conjoint Analysis (CA). CA is an individual analysis based on the observed evaluation behavior of one specific individual [35]. The observed behavior is used to define a preference, which is a one-dimensional indicator of individual's preference structure [36]. The structure describes what object is favored by the individual. While compositional methods ask individuals about their preference for certain attributes and compose an overall judgment from it, decomposition methods, such as the CA, calculate the partial utility values for each attribute from the overall judgement of the participants [36].

3 Empirical Study

3.1 Methodological Approach

In the current study the choice-based conjoint analysis (CBC) was chosen because of its methodological and practical strengths [37]. CBC is based on the work of Louviere and Woodworth from 1983 [38] and combines the discrete choice analysis (DCA) with the Traditional Conjoint Analysis (TCA) [37]. Therefore, it is measuring population's utility functions. Those functions are estimated by representative utility functions. First, the most important assumptions are that participants always choose the product profile with the highest individual utility. Thus, it is possible to draw conclusions from the purchase decisions and the utility functions of the users [37]. Second, it is assumed that the utility function consists of a deterministic and a stochastic component, which are summed up. The main difference between CBC and TCA is that instead of ranking stimuli, the CBC wants its participants to rank different choice tasks. Those options consist of a product with a bundle of chosen attributes and their levels [37]. Instead of ranking different profiles against each other like TCA, participants have to perform fictitious purchase decisions [39]. Those choice tasks consist of the predetermined product profiles (choices) that display the attributes and their levels [37]. In the current paper a CBC following the steps of Backhaus et al. [40] was designed: definition of stimuli, design choice situation, utility model, choice model, and estimation of utility values. The CBC proves to avoid the distortions in surveys caused by group dynamics and social desirability. The advantages of the indirect measurement of preferences for certain product attributes utilizing the CBC approach shine especially against the background of the privacy paradox.

3.2 Survey Design

Apps where chosen as research objects, due to their broad diffusion in mass user markets and their everyday life integration. To ensure participants common understanding regarding e.g. functionality, provider, and the privacy level of access privileges an appealing CampusApp was defined and conducted at a German university. The functionality of the app was designed similar to campus apps of comparable universities (navigation on campus, information about public transportation, library services, organization of studies incl. online platform of the university, food on campus, university sports programs).

To keep a low stress level for the study participants, the number of attributes was set to four. To determine the attributes, different steps were conducted. First, the recent literature about apps' product attributes, as well as their assigned categories and their influence to the users' decision-making process when purchasing an app were analyzed [41]. Second, a word-frequency text analysis of 73 apps from the categories 'most popular' and 'top 10' apps with and without a purchase price was performed. In the third step the available types of information in the two most common app stores (Android Play Store and Apple App Store) were examined. Their separate information was compared to one another, as well as categorized into five different groups. With the broad variety of app product attributes, an online survey (N=151) to estimate the perceived importance of the different information types was conducted and narrowed down to a suitable amount of four attributes for the CBC. Attributes and levels are shown in table 1.

Table 1. App Attributes and Levels by Groups

Groups	Level 1	Level 2	Level 3
I - Price	0,00 €	0,99 €	2,77 €
II - Privacy	Only functionally required permissions requested with privacy policy	More than functionally required permissions requested with privacy policy	More than functionally required permissions requested without privacy policy
III - Rating	4 stars	3 stars	2 stars
IV - App	0-500 MB	500 MB - 1 GB	> 1 GB

The attribute's levels of the price group (I) were defined by taking a closer look at the common prices in the app stores. With over 60% of all apps since 2009, the price of 0,00 € is by far the most common [42]. Although the average app price is reported between $1.13 [43] and $1.91 [44] by different sources, researchers agree that this price is decreasing. Following, and to design an attractive price in the middle for the second level, 0,99€ was determined. In order to create a realistic high-end price, the price of 2,77€ was chosen. This price is based on a bidding game for a messenger app by Buck [45].

In the privacy-related group (II), the handling of personal user data and the technical access to personal user data were combined into one attribute. Taking a closer look at the current handling of personal user data, Sunyaev, Dehling, Taylor, and Mandl's [46] work shows that only 30.5% of the examined mobile health apps had a privacy policy. Additionally, and since advertising is one of the most used mobile app monetization models of developers [47], apps frequently request more technical access and permissions than they actually need to function properly. Based on those findings, only the first level of privacy was designed to request the required amount of permissions and a privacy policy. The other two levels both requested more technical permissions than necessary. The difference between those last two levels is that the second level possesses a privacy policy, but the third level does not. In the survey itself, required permissions were displayed with an exemplary set of functionally required permissions and an exemplary set of permissions which exceed the functionally required amount and permission types significantly. Both sets were modeled on the basis of permission groups given in the Android OS, which are very similar to the ones in the iOS.

Within the app ratings group (III), the average rating in stars was the only attribute reaching a Likert-scale average above five. Although rankings and ratings in the app stores have proven to suffer from fraud [48], based on the pre-study, app users are familiar and relying on the star ratings. Following, the levels for average app rating in stars were determined to two, three, and four stars. Zero stars and five stars were not considered because those ratings mostly consist of as little as one review or none at all.

In the last considered group of directly app-related information attributes (IV), the 'compatibility with own devices' was the most important attribute. Due to the fact that this is certainly a deal-breaker attribute for users [35], the second most important attribute was chosen for that group: needed size on device in MB/GB. Since there are all kinds of different apps, the levels were set to <500 MB, 500 MB -1 GB and >1 GB.

To avoid the phenomenon of forced choice and to design the choice situation as realistic as possible, an additional 'none' option is included [40]. Additionally, an unrealistic set containing level one of price and privacy were not included to ensure a realistic choice situation for the participants. With the four chosen attributes their three levels, a total of 81 different stimuli sets are possible. The stimuli were presented in the form of virtual cards with descriptions.

In order to match the recommended range, a fractional design with 10 randomized stimuli and two fixed stimuli was chosen. Based on the recommended number of choices per set of $K \leq 7$ [40], four randomized choices and a 'none' option were chosen for the CBC. The fixed choice tasks were designed to confront the participants with a trade-off situation between price and privacy, as shown in figure 1.

In choice 1 and 3, the attributes of rating and app-related are both marked as level 1. The difference between those two options is to be found in the attributes of price and privacy. Choice 1 has a level 1 privacy scheme, but a level 3 price (2,77€). Choice 3 has a level 1 price (0,00€), but a level 3 privacy scheme. Choice 2 and 4 are designed to be middle-class options with level 2 compositions for price and privacy. Nevertheless, choice 2 is dominated by choice 4, since it is shaped significantly worse in the rating and size attributes. With four attributes and three levels each in a choice situation with a choice set of four plus a 'none' option, so-called overlaps occur in every single choice set. These overlaps allow improvement in the measurement of precise interactions between the attributes [49]. Still, and in order to prevent impacts that are too drastic on the main effects, the balanced overlap was chosen as task generation method [49]. Since the recommended range of choice tasks is 8 to 20 and the participants' concentration decreases significantly with every choice task [50], only 10 randomized choice tasks were displayed.

Figure 1. Exemplary Random Choice Set with Four Stimuli

3.3 Data Collection

The CBC was conducted via an online survey using Sawtooth Software. Since the study was conducted at a German university, all questions were presented in German only. At the beginning of the study, a skip logic question was given with the aim to select only participants with ties to this specific university. This was established to ensure a minimum involvement regarding the usage and functionalities were given and the incentive for downloading the CampusApp was comparable. Following, the participants were asked technical context information (e.g. SMD usage, on-device installed apps, app downloading habits, and buying likelihood of apps in the near future). Afterwards, the participants were asked to rank six different attributes according to their importance when considering to download an app. In order to introduce the CBC, the CampusApp was explained with an image and a list of functions. A second explanation including how the following choice sets will look and introductions for the CBC were displayed on next screen.

As outlined before, the CBC itself consists of 10 randomized, as well as two fixed, choice tasks. The fixed choice tasks were designed to examine a direct trade-off between price and privacy, as well as to conduct a hold-out analysis to predict the prognosis validity. After the CBC, participants were asked for their gender, age, and the brand of their SMD.

In total, 221 respondents participated in the online survey. However, 71 responses were incomplete and therefore excluded from the analysis. The participants who reported having no existing relationship to the university in the skip logic question were part of this exclusion as well. Additionally, all participants who answered more than 50% 'none' in the choice tasks were also eliminated. This step was conducted in order to reduce the weakening impact of the attribute utility values and their levels. In total, the results of 111 responses were analyzed.

3.4 Results

Out of the 111 participants who partook in the survey, 43% were female and 57% were male. The average age (mean value) of all participants was 24.99 years. The participants were asked to rank six different attributes by importance when purchasing an app (Table 2). In order to prevent the participants from focusing on any one specific attribute and eventually influencing their answering behavior later, two additional attributes were included in the conscious priority ranking: Vendor's Reputation and Number of Ratings.

Table 2. Priority Ranking

Attribute	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Reputation	8,11%	29,73%	0,90%	48,65%	9,01%	3,60%
Av. Rating	10,81%	36,04%	15,32%	16,22%	16,22%	5,41%
No. of Ratings	12,61%	15,32%	24,32%	16,22%	19,82%	11,71%
Price	17,12%	14,41%	20,72%	10,81%	23,42%	13,51%
Privacy	30,63%	2,70%	18,92%	7,21%	23,42%	17,12%
Space	20,72%	1,80%	19,82%	0,90%	8,11%	48,65%

Ranked with just over 30%, most participants named privacy as rank 1. Privacy is followed on rank 1 by required space (MB/GB) with 20.72% and price with 17.12%. In rank 2, average rating with 36.04%, vendor's reputation with 29.73%, and number of ratings with 15.32% was valued most. Based on those consciously rated importance rankings, the first indication for privacy's crucial status among app product attributes occurs.

Based on the choices made by the participants in the CBC section of the online survey, average utilities for each level of each attribute were calculated by using Sawtooth Software. As for analysis type, Hierarchical Bayes (HB), the go-to standard for utility estimations in CBC, was chosen. In total, a number of 20,000 iterations were used. Hereby, only the second 10,000 iterations were used to avoid assuming convergence too early.

As a result, utility values and standard deviations for all attributes' levels, as well as the 'none' option, are calculated (Table 3). At first sight, the negative impacts of the third levels of each attribute are noticed. In contrast to that, the first and second levels of all attributes have a positive impact on the individuals' utility values. Nevertheless, it is not possible to tell how much more or less importance an attribute or its levels have while solely regarding absolute utility values. Therefore, the average importance of each attribute is calculated in percentages based on the relative utility ranges (Table 4).

Table 3. Average Utilities (Zero-Centered Diffs)

Attributes' Levels	Average Utilities	SD
4 stars	52.25	32.13
3 stars	8.75	8.63
2 stars	-61.00	29.65
Permissions (functional) & privacy policy	52.11	46.28
Permissions (more) & privacy policy	16.25	18.83
Permissions (more) & no privacy policy	-68.36	41.65
0 - 500 MB	6.23	14.63
500 MB - 1 GB	6.57	11.90
> 1 GB	-12.80	12.04
0,00 €	53.36	34.80
0,99 €	10.40	14.44
2,77 €	-63.76	33.27
NONE	24.67	47.74

Taking a closer look at the 'Average Importance', privacy turns out to be the most important attribute with over 32%. Closely after privacy follows price with just over 30%. On the third position sits average rating with nearly 29%. With around 8%, required space (in MB/GB) is least important.

Table 4. Average Importance

Attributes	Average Importance	SD
Average Rating	28.92	14.27
Permissions & Privacy Policy	32.80	18.38
Required Space (MB/GB)	8.06	4.40
Price	30.21	15.41

Comparing the results of the consciously ranked attributes and the CBC-based results of the HB estimation for attribute importance, privacy is named as the most important attribute when buying an app in both cases. Privacy is reported as rank 1 priority with over 30% in the direct ranking question, as well as calculated as most important through the CBC's utility estimations.

Although required space is named second-important in the priority ranking, the importance percentage of only 8% shows that users do not actually value this attribute as much as they state. Price, with 17.12% is third-ranked in the priority ranking. The importance of this specific attribute is validated by the results of the CBC's HB estimation. With over 30%, price's importance is second-placed. Additionally, average rating is ranked most important on rank 2 in the priority rating. The importance, slightly below privacy and price, is to be found in the percentages of the HB estimation as well.

Tests show high values for face validity, intern validity, and prognosis validity [51]. The hit rate of 68.47% indicates decent results for the study. The study's average root likelihood is 0.6 which proves an accurate internal validity. The hit rate of 76.58% is significantly bigger than 20% which shows a promising prognosis of validity for the study.

4 Privacy as a Crucial Product Attribute

Concluding, the results of both examinations of the attributes' importance through the priority ranking, as well as the HB-based utility value and importance estimation of the CBC, show significantly high levels of importance for privacy. In fact, privacy is on first-place in both rankings. Following the research question of whether or not privacy in the area of SMD and SMA represents a crucial product attribute for the user, must be affirmed. The fact that privacy ranks even more important than price in the consciously answered priority ranking, as well as in the CBC, shows an exceptional observation. Especially the result that privacy is ranked as no. 1 in the preference structure provides novel insights in users' intention when downloading apps. In contrast to many WTP-studies, where users were only willing to pay a very small amount of money for their privacy, the results of the conducted CBC suggest that there is a high preference for controlling privacy. This could indicate that users are willing to pay a higher purchase price than they currently do, when their privacy protection is ensured and promoted as an outlined product attribute.

In consequence, the results show valuable implication for theory and practice. The significantly high level of privacy in importance, which is even higher than price, indicates that SMD users demand more options to handle their user data and to protect their privacy. In contrary, customers of the two biggest app stores do not usually have the option of choosing between paying a monetary price or revealing their private user data. This imbalance provides a huge potential for innovating apps business models and its monetization. As for now, the user mostly has to decide if the apps provided utility is worth a privacy intrusion or not – meaning the user cannot use the app although a certain willingness-to-pay might exist. The study shows that at least offering apps in an alternative version with a monetary price and no usage of private user data could bear a great potential for success. Other options, such as permission management or administration of user data, could be another potential, but would represent a more restrictive way to deal with the privacy issue within SMD and apps. Since most of the apps requesting permissions regarding private user data do not function correctly without certain permissions, this approach might prove difficult to provide the full amount of utility of an app while containing only restricted permissions. Nevertheless, first developments to single permission management for each app in the Android OS are observable when taking a closer look at the newest OS 'Android 6.0 Marshmallow' [52].

5 Limitations and Future Research

Our paper deals with the question whether privacy is a crucial product attribute for users when buying apps. The results of the conducted CBC outline privacy as the product attribute ranking at the highest importance level and generating the highest utility value.

Due to the nature of our research, our study has some limitations. For example, in this paper we refer to the ‘download’ or purchase of apps. However, we are aware, that disclosing personal data is also related by app usage and deletion, which should be considered in future studies related to the topic. Furthermore, a particular app as a study object was required. Since the variety of apps could not be displayed with one app and the functionality has to be defined for the study object, the fictional CampusApp was selected. Our sample is not representative of all app users, as it includes a large group of university related participants. Based on the choice of the study object mostly students and employees of the addressed university were asked to answer the CBC. Although the focus on the four product attribute groups of price, privacy, ranking, and app-related was necessary due to the CBC complexity and justified by the low importance of vendor-related attributes in the pre-study, the results indicate that at least in the conscious priority ranking, vendor’s reputation was considered quite important. Moreover, the preference structures’ stability is questionable over time.

Starting with the high importance level of privacy within the purchase situation of apps, other privacy-sensitive areas, like private banking, insurance services, online social networks, or all kind of digital services linked with personal data could be investigated in more detail by using the CBC. This leads to the need of contemporarily and repeatedly conducted CBC in the future to maintain the topicality of the results and to validate privacy’s standing as a crucial product attribute. Future CBC in the area of SMD and apps might include vendor-related attributes. Moreover, the elimination of choice sets could distort the results. Especially in the light of low effort situations and behavioral effects [25], taking a closer look at apps from various categories might offer interesting insights. Based on the social, political, legal, and additionally personal salience of privacy, further research in the area of privacy is essential. For example, the understanding and determination of the value term regarding privacy, the valuation of users’ preference structures, or the explanation of the privacy paradox, offer a great deal of opportunities for future theoretical research and for a deeper understanding for more adequate attempts to assign a monetary value to privacy. Practically, research might focus on the options of permission management without losing functionality or the economical options of offering the same app in different versions regarding price and privacy. Additionally, those implications for apps, as well as SMD, and their relation towards privacy are also applicable to a very wide range of different research topics, e.g. investigation of the influence of privacy on different app types, as well as different demographical, social, or national groups.

Following the understanding of privacy as a crucial product attribute of apps, the legal regulation is called upon to preserve users of disclose their personal data in purchase situation they cannot control.

References

1. Weiser, M.: The computer for the 21st century. Scientific american 265, 94–104 (1991)
2. Yoo, Y.: Computing in everyday life: A call for research on experiential computing. MIS quarterly, 213–231 (2010)
3. Venkatesh, V., Thong, J.Y.L., Xu, X.: Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS quarterly 36, 157–178 (2012)
4. Smith, H.J., Dinev, T., Xu, H.: Information Privacy Research: An Interdisciplinary Review. MIS quarterly 35, 989–1016 (2011)
5. Solove, D.J.: A taxonomy of privacy. University of Pennsylvania Law Review, 477–564 (2006)
6. Westin, A.F.: Social and political dimensions of privacy. Journal of social issues 59, 431–453 (2003)
7. Westin, A.F.: Privacy and Freedom, Atheneum. New York, 7 (1967)
8. Wellman, B.: The reconstruction of space and time: Mobile communication practices. Contemporary Sociology: A Journal of Reviews 39, 179–181 (2010)
9. Buck, C., Horbel, C., Kessler, T., Germelmann, C.C.: Mobile consumer apps: big data brother is watching you. Marketing Review St. Gallen 31, 26 (2014)
10. Clarke, R.: Internet privacy concerns confirm the case for intervention. Communications of the ACM 42, 60–67 (1999)
11. Chen, H.-T., Chen, W.: Couldn't or wouldn't? The influence of privacy concerns and self-efficacy in privacy management on privacy protection. Cyberpsychology, Behavior, and Social Networking 18, 13–19 (2015)
12. Dinev, T., Hart, P.: An extended privacy calculus model for e-commerce transactions. Information Systems Research 17, 61–80 (2006)
13. Acquisti, A., Brandimarte, L., Loewenstein, G.: Privacy and human behavior in the age of information. Science 347, 509–514 (2015)
14. Word Economic Forum: Personal Data: The Emergence of a New Asset Class, http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf
15. Spiekermann, S., Acquisti, A., Böhme, R., Hui, K.-L.: The challenges of personal data markets and privacy. Electron Markets 25, 161–167 (2015)
16. Bennett, C.J.: The political economy of privacy: a review of the literature. center for social and legal research, DOE genome project (Final draft), University of Victoria, Department of Political Science, Victoria (1995)
17. Berthold, S., Böhme, R.: Valuating Privacy with Option Pricing Theory. In: Economics of information security and privacy, pp. 187–209
18. Flender, C., Müller, G.: Type indeterminacy in privacy decisions: the privacy paradox revisited. In: Quantum Interaction, pp. 148–159. Springer (2012)
19. Buck, C., Germelmann, C.C., Eymann, T.: Datenweitergabe als Bedrohung? Konsumentenwahrnehmung am Beispiel mobiler Applikationen. In: Schmidt-Kessel, M., Langhanke, C. (eds.) Datenschutz als Verbraucherschutz, pp. 49–67. JWV Jenaer Wissenschaftliche Verlagsgesellschaft, Jena (2016)
20. Culnan, M.J., Armstrong, P.K.: Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. Organization science 10, 104–115 (1999)
21. Malhotra, N.K., Kim, S.S., Agarwal, J.: Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. Information Systems Research 15, 336–355 (2004)

22. Xu, H., Teo, H.-H., Tan, Bernard C. Y., Agarwal, R.: The role of push-pull technology in privacy calculus: the case of location-based services. Journal of Management Information Systems 26, 135–174 (2009)
23. Pavlou, P.A., Gefen, D.: Building effective online marketplaces with institution-based trust. Information Systems Research 15, 37–59 (2004)
24. Norberg, P.A., Horne, D.R.: Privacy attitudes and privacy-related behavior. Psychology & Marketing 24, 829–847 (2007)
25. Dinev, T., McConnell, A.R., Smith, H.J.: Research Commentary: Informing Privacy Research Through Information Systems, Psychology, and Behavioral Economics: Thinking Outside the "APCO" Box. Information Systems Research 26, 639–655 (2015)
26. Schreiner, M., Hess, T.: On The Willingness To Pay For Privacy As A Freemium Model: First Empirical Evidence. In: ECIS 2013 (2013)
27. Chellappa, R.K., Sin, R.G.: Personalization versus privacy: An empirical examination of the online consumer's dilemma. Information Technology and Management 6, 181–202 (2005)
28. Wathieu, L., Friedman, A.A.: An Empirical Approach to Understanding Privacy Valuation. SSRN Journal (2007)
29. Spiekermann, S., Korunovska, J., Bauer, C.: Psychology of Ownership and Asset Defense. Why People Value Their Personal Information Beyond Privacy. SSRN Journal (2012)
30. Bauer, C., Korunovska, J., Spiekermann, S.: On the value of information - what facebook users are willing to pay. ECIS 2012 Proceedings (2012)
31. Krasnova, H., Eling, N., Abramova, O., Buxmann, P.: Dangers of 'Facebook Login' for Mobile Apps: Is There a Price Tag for Social Information? ECIS 2014 Proceedings (2014)
32. Hann, I.-H., Hui, K.-L., Lee, T., Png, I.: Online information privacy: Measuring the cost-benefit trade-off. ICIS 2002 Proceedings, 1–10 (2002)
33. Krasnova, H., Hildebrand, T., Guenther, O.: Investigating the value of privacy in online social networks: conjoint analysis (2009)
34. Pu, Y., Grossklags, J.: Using Conjoint Analysis to Investigate the Value of Interdependent Privacy in Social App Adoption Scenarios. ICIS 2015 proceedings (2015)
35. Böhler, H., Scigliano, D.: Traditionelle Conjointanalyse (in German). In: Baier, D., Brusch, M. (eds.) Conjointanalyse: Methoden, Anwendungen, Praxisbeispiele (in German), pp. 101–112. Springer, Berlin, Heidelberg (2009)
36. Baier, D., Brusch, M.: Erfassung von Kundenpräferenzen für Produkte und Dienstleistungen (in German). In: Baier, D., Brusch, M. (eds.) Conjointanalyse: Methoden, Anwendungen, Praxisbeispiele (in German), pp. 3–19. Springer, Berlin, Heidelberg (2009)
37. Balderjahn, I., Hedergott, D., Peyer, M.: Choice-Based Conjointanalyse (in German). In: Baier, D., Brusch, M. (eds.) Conjointanalyse: Methoden, Anwendungen, Praxisbeispiele (in German), pp. 129–146. Springer, Berlin, Heidelberg (2009)
38. Louviere, J.J., Woodworth, G.: Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. Journal of marketing research, 350–367 (1983)
39. Cohen, S.H.: Perfect Union. CBCA marries the best of conjoint and discrete choice models. Marketing Research, 12–17 (1997)
40. Backhaus, K., Erichson, B., Weiber, R.: Fortgeschrittene Multivariante Analysemethoden (in German). Eine anwendungsorientierte Einführung, Berlin (2011)
41. Buck, C., Horbel, C., Germelmann, C.C., Eymann, T.: The Unconscious App Consumer: Discovering and Comparing the Information-Seeking Patterns among Mobile Application Consumers. ECIS 2014 Proceedings (2014)
42. Statista: Average prices for apps in the Apple App Store as of January 2016 (in U.S. dollars), http://www.statista.com/statistics/267346/average-apple-app-store-price-app/

43. Cowley, R., Suckley, M. and Jordan, J.: App Store Metrics, http://www.pocketgamer.biz/metrics/app-store/app-prices/
44. Statista: Schätzung des durchschnittlichen Preises kostenpflichtiger Apps für das iPhone und iPad weltweit in den Jahren 2009 bis 2022 (in US-Dollar), http://de.statista.com/statistik/daten/studie/170003/umfrage/preisentwicklung-von-apps-in-den-fuehrenden-app-stores-weltweit/
45. Buck, C.: App-privacy as an abstract value – Approaching contingency valuation for investigating the willingness to pay for app-privacy (2015)
46. Sunyaev, A., Dehling, T., Taylor, P.L., Mandl, K.D.: Availability and quality of mobile health app privacy policies. Journal of the American Medical Informatics Association (2014)
47. Statista: Mobile app monetization (2015)
48. Zhu, H., Xiong, H., Ge, Y., Chen, E.: Ranking fraud detection for mobile apps. In: He, Q., Iyengar, A., Nejdl, W., Pei, J., Rastogi, R. (eds.) the 22nd ACM international conference, pp. 619–628 (2013)
49. Sawtooth Software Inc.: The CBC System for Choice-Based Conjoint Analysis. Version 8
50. Johnson, J., Huber, J., Orme, B.: A Second Test of Adaptive Choice Based Conjoint Analysis. (The Surprising Robustness of Standard CBC Designs). In: Sawtooth Software Inc. (ed.) Proceedings of the Sawtooth Software Conference on Perceptual Mapping, Conjoint Analysis and Computer Interviewing, pp. 219–236 (2004)
51. Gensler, S.: Ermittlung von Präferenzen für Produkteigenschaften mit Hilfe der Choice-Based Conjoint Analyse, Teil II. Frankfurt am Main (2006)
52. Google Inc.: Android 6.0 Marshmallow, https://www.android.com/intl/de_de/versions/marshmallow-6-0/

Beyond Mere Compliance – Delighting Customers by Implementing Data Privacy Measures?

Dominikus Kleindienst[1], Niclas Nüske[1], Daniel Rau[1], and Fabian Schmied[1]

[1] FIM Research Center, University of Augsburg, Augsburg, Germany
{dominikus.kleindienst,niclas.nueske,daniel.rau,fabian.schmied}@fim-rc.de

Abstract. The importance of customer data for business models is increasing, as is the relevance of customers' concerns regarding privacy aspects. To prevent data privacy incidents and to mitigate the associated risks, companies need to implement appropriate measures. Furthermore, it is unclear whether their implementation – beyond mere compliance – has the potential to actually delight customers and yields competitive advantages. In this paper, we derive specific measures to deal with customers' data privacy concerns based on the literature, legislative texts, and expert interviews. Next, we leverage the Kano model via an Internet-based survey to analyze the measures' evaluation by customers. As a result, most measures are considered basic needs of must-be quality. Their implementation is obligatory and is not rewarded by customers. However, delighters of attractive quality do exist and have the potential to create a competitive advantage.

Keywords: Data Privacy, Customer Satisfaction, Kano Model

1 Introduction

With the growing amount of data generated worldwide, digital business models emerge that are based on insights gained from customer data [1-2]. At the same time, trust in data privacy is becoming more relevant for customers [3-4], which is amplified by several data privacy scandals in the recent past. Examples range from Ashley Madison, an online dating portal that lost user data of 37 million registered married men and women to the public, to Apple, which was accused of collecting location data on iPhones and iPads without authorization from and without notifying their customers, to Facebook, which was discovered to be collecting data from user profiles and transmitting these data to advertising companies and others. For companies, such publicly exploited scandals cause economic damage [5-6] and competitive disadvantages in brand image and customer satisfaction. Thus, companies that perform well in improving data privacy could increase customer satisfaction and gain a competitive advantage. For instance, companies such as DuckDuckGo or Silent Circle already try to differentiate themselves by providing privacy friendly services [7]. However, for many companies, it is often unclear how to manage data privacy, which is viewed as a necessary evil. As such, data privacy limits the opportunities to gain

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Kleindienst, D.; Nüske, N.; Rau, D.; Schmied, F. (2017): Beyond Mere Compliance — Delighting Customers by Implementing Data Privacy Measures?, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 807-821

valuable customer insights, and its implementation binds valuable resources. In addition to that downside perspective, for integrated management of data privacy, an upside perspective is also necessary. Moreover, practitioners should be aware of specific available data privacy measures that enable their companies to differentiate themselves from their competitors.

In the literature, data privacy management is mostly seen from a downside perspective that focuses on risk management. For instance, Buhl [8] state that data privacy measures should be implemented only if the risk-reducing effects outweigh the related costs. Acquisti et al. [5] link a company's privacy incidents to the negative impacts on its market value. Only to a small extent does the literature consider an upside perspective on data privacy, such as Preibusch et al. [4], who found that customers of privacy-friendly but more expensive firms are more satisfied than customers of cheaper but privacy-unfriendly firms. Even so, specific data privacy measures that might be implemented to increase customer satisfaction are yet to be considered in the literature. Thus, we investigate the following research question: Can companies delight customers by implementing specific data privacy measures?

To answer this research question, we firstly develop an overview of data privacy measures by investigating and consolidating the literature, legislative texts, and findings from expert interviews. Secondly, using the Kano model, we evaluate customers' perception of these different measures, that is, whether different measures are considered "must-be," "one-dimensional," or "attractive," or whether customers are "indifferent." Thereby, this paper is organized as follows. We discuss the context of the problem and related work. Using this discussion, we outline our methodical approach, derive measures that can be taken by companies to address data privacy concerns, and analyze customers' perceptions of these measures on the basis of the results of a survey. The conclusion summarizes the results, addresses limitations, and discusses areas of possible future research.

2 Problem Context

As previously motivated, public attention to data privacy issues is growing. This attention is reflected in different scientific disciplines, such as philosophy, psychology, economics, marketing, law, and information systems [9-10]. Moreover, privacy incidents, such as the scandals previously mentioned, are the subject of research projects (e.g., [5], [11-14]). Privacy incidents appear regularly and have consequences for both companies and customers. They are defined by Acquisti et al. [5] as events "involving misuses of individuals' personal information." Consequently, customers might become victims of fraud or identity theft [5]. Typical customers' data privacy concerns are composed by Smith et al. [15], namely, Data Collection or Combination, Internal and External Secondary Usage, Errors, Improper Access, and Reduced Judgement. The literature provides recommendations for customers and public authorities responsible for protecting customers' privacy rights through laws and regulations [16-17]. From a company perspective, privacy incidents may be caused by technical, managerial, organizational, or human failures [5]. Companies might suffer direct economic damage, such as punishment by penalties or loss of market value, as well as indirect effects, such as increasing insurance fees or decreasing customer satisfaction [5], [11].

Consequently, companies must decide on how to deal with data privacy issues and the related risks. In line with that issue, articles that address companies' handling of data privacy focus on potential threats and how to avoid their occurrence. Conversely, only a limited set of articles considers data privacy measures as an opportunity to create a competitive advantage. For instance, Preibusch et al. [4] show that appropriate management of data privacy issues may have positive implications on customer satisfaction, whereas Sarathy and Robertson [18] provide a framework that assists companies in implementing a data privacy strategy that considers ethical aspects. However, neither article provides recommendations for specific data privacy measures that can be implemented to address customers' data privacy concerns and increase customer satisfaction.

Hence, to the best of our knowledge, using data privacy to delight customers to gain a competitive advantage has yet to be comprehensively examined. More precisely, the literature has yet to provide insights into addressing customers' different privacy concerns using concrete measures and the extent to which such measures affect customer satisfaction. Thus, we raise the research question of whether companies can delight customers by implementing specific data privacy measures.

3 Research Method

To answer the research question, we firstly need to structure the field of possible data privacy measures. Accordingly, this section starts with outlining the identification process of possible data privacy measures. After a short discussion of models that intend to measure customer satisfaction, we describe how we used the Kano model to evaluate customers' perceptions regarding different data privacy measures. The third part of this section describes the design and participants of an online survey used to collect customers' evaluations.

3.1 Identification of Data Privacy Measures

As a basis for identifying data privacy measures, we conducted a comprehensive search for relevant statements, that is, any piece of information on any type of action that addresses customers' data privacy concerns. Therefore, our sources are legislative texts in particular (European General Data Protection Regulation, Bundesdatenschutzgesetz, Telemediengesetz), but also scientific and practitioner-oriented literature found in the databases Springerlink, AISeL, ProQuest, ScienceDirect, EBSCO host, and JSTOR without timeframe limitation and including back-forward search. However, we find that literature does only address data privacy measures to a very small extent. Thus, despite having found 31 relevant papers dealing with managing data privacy in terms of data privacy measures, only two papers contain statements relevant for deriving feasible data privacy measures. Additionally, we conducted three expert interviews, each lasting approximately 30 to 60 minutes and being divided in two parts. The first part was a free talk with the goal of gaining new insights. In particular, this part was used to identify additional statements regarding data privacy measures. In the second part, already identified statements from other sources were evaluated by the interviewee. All statements were then grouped by semantic similarity. In doing so, all authors jointly decided on the grouping of the statements. Without having pre-defined groups, each statement was either used to create a new group with a particular data privacy measure or mapped to an existing group. As a result, all groups consisted of one or several statements regarding a particular data privacy measure. From each of the groups of statements, we derived a single measure that addresses all statements within the group. After the formulation of the measures, we assigned each of the measures to one or more specific customer data privacy concerns.

3.2 Kano Model

After the derivation of data privacy measures and their assignment to specific data privacy concerns, we now focus on determining their effect on customer satisfaction. Approaches to measuring customer satisfaction can be differentiated into subjective and objective methods. The latter can be further distinguished between event-oriented, problem-oriented, and attribute-oriented techniques [19]. As our research focusses on service attributes (measures), we focus on attribute-oriented techniques. In this context, the most commonly used method to measure service quality is SERVQUAL [20-21], but also structural equation modelling and neural networks [22] are viable options. Bartikowski and Llosa [23] analyze further methods, namely Penalty Reward Contrast Analysis, Correspondence Analysis, Dual Importance Mapping, and the Simulation Method. The context at hand requires the possibility of individual investigation of each measure and applicability to hypothetical cases. These criteria are only fulfilled by the Simulation Method, which is most prominently represented by Kano [23]. Accordingly, we use the Kano model to determine customers' evaluation of the identified data privacy measures. The Kano model has been discussed and applied in many theoretical and empirical research projects [24-25], as it provides a comprehensible presentation of attributes of products or services which influence the degree of customer satisfaction. For instance, the model has been used by Lai and Wu [26] in order to gain insights in the customers' needs of a Taiwanese public transport company and by Arbore and Busacca [27], who studied determinants of customer satisfaction for an Italian retail bank.

The Kano model describes customer satisfaction on the basis of the degree of implementation or availability, respectively, of certain attributes of products or services [28-29]. The model differentiates between four major types of factors. In Table 1, we list the factors and apply the remarks of Matzler et al. [28] to data privacy measures. In Figure 1, we illustrate the dimensions and possible factors of the Kano model in the context of data privacy measures that refer to Matzler et al. [28].

Table 1. Details on the factors of the Kano model [29] as described by Matzler et al. [28]

Factor	*Customers' expectations and resulting effect on satisfaction*
Delighter (attractive quality)	Customers do not expect implementation of measure: • Implementation has a positive effect on satisfaction • Non-implementation has no effect on satisfaction
Performance need (one-dimensional quality)	Customers explicitly demand implementation of measure: • Implementation has a positive effect on satisfaction • Non-implementation has a negative effect on satisfaction
Basic need (must-be quality)	Customers implicitly demand implementation of measure: • Implementation has no effect on satisfaction • Non-implementation has a negative effect on satisfaction
Indifferent quality	Customers are indifferent to implementation of measure: • Implementation has no effect on satisfaction • Non-implementation has no effect on satisfaction

To determine the categorization of customer requirements as one of the Kano model factors, it is most common to use a two-question approach [25]. This original approach by Kano has been found to be the most reliable one in a comparison with four other methods and to be one of only two approaches that are suitable to be used in the design stage of products or services [30].

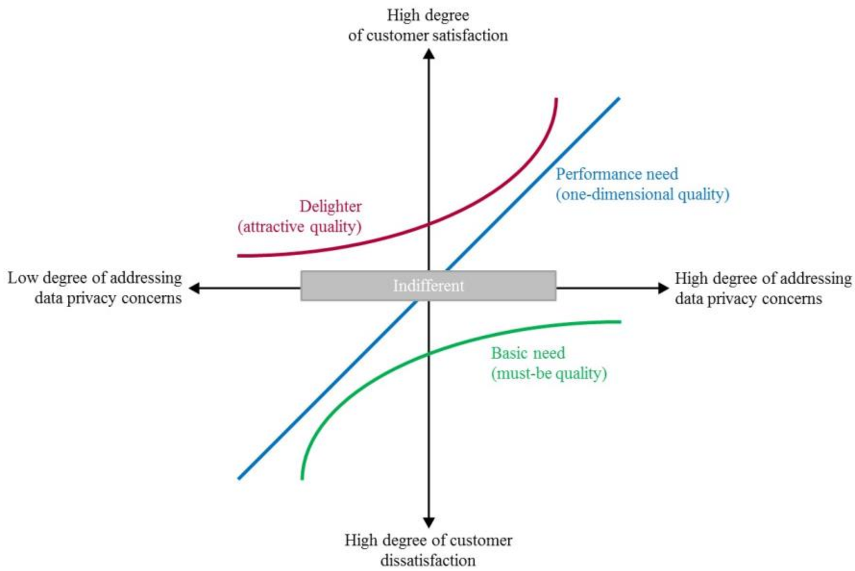


Figure 1. Factors of the Kano model [29] as described by Matzler et al. [28] and applied to the context of data privacy

The classification of a measure as a certain factor depends on customers' answers to both a functional and a dysfunctional question. That is, customers are asked about their evaluation of the hypothetical case in which a measure is implemented and a case in which it is not. Each time, they can choose one of five possible answers: “I like it that way,” “It must be that way,” “I am neutral,” “I can live with it that way,” and “I dislike it that way.” The different answers do not stand for a level of acceptance and there is no ordinal scale. According to Kano et al. [29], each possible combination of answers can be interpreted in an individual manner and leads to a certain pre-defined classification [28], as shown in Figure 2. As proposed by Matzler et al. [28], we derive the final classification of a measure from the respective most frequent individual result.

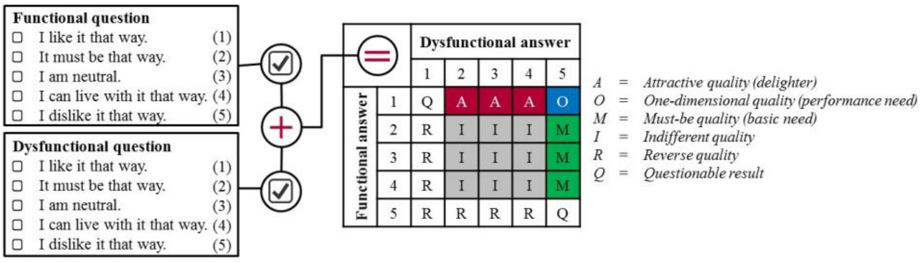


Figure 2. Derivation of Kano model factors based on Matzler et al. [28]

3.3 Survey

Scenario. In order to determine the customers’ evaluation of the identified data privacy measures, we conduct an Internet-based survey. To enable the participants to assume a perspective as natural as possible and to illustrate the situation, we need to use a specific, well-known, and simple scenario that relates to an exemplary industry sector, for which data privacy is a considerable issue. That is, the sector should feature a business-to-consumer market with a significant occurrence of processed customer data. To be able to consider the possible exchange of customer data between companies, cooperation agreements should exist between major industry actors. Furthermore, companies should provide loyalty programs because they are typically based on gathering data on a customer’s behavior over a long period.

The aviation sector as a commonly known industry with a considerable amount of customer data collected at different interaction points [31], transmission of data to public authorities, airport operators, or other airlines that are partners in global alliances [32], and loyalty programs, fulfills all of these requirements.

Design. To ensure high quality results, we first ran a pretest followed by the main survey. In the pretest, we asked 85 German-speaking participants to imagine booking a flight through an airline’s website. Each participant was asked a functional and a dysfunctional question for each of the measures. Using the insights of the pretest, we made several modifications to the main survey: for improving the response rate, we mixed the questions with invitations to guess the correct answers to fun-fact questions about the aviation sector. For improving understandability, we grouped the questions with regard to data privacy concerns preceded by short explanations of the respective concerns. The following example of an explanation, a functional, and a dysfunctional question demonstrates the survey's design. *Explanation:* "Your customer data may be used by a third party outside of the company for a purpose not previously agreed upon. The company implements the following measures." *Functional question:* "You are informed if your customer data are passed on to external third parties." *Dysfunctional question:* "You are not informed if your customer data are passed on to external third parties." To answer the functional and the dysfunctional question, the participants can choose one of the five previously mentioned possible answers. In this way, we ask the participants about each of the 32 identified measures, resulting in a total of 32 question pairs, each of them addressing one of the data privacy concerns.

Participants. The main survey has 227 German-speaking participants, 219 of whom correctly answered a control question. Invitations were distributed via social media and email, and participation was incentivized through a lottery of vouchers for an online retailer. The sample mostly consists of students (78%) and employees (16%). The age of the participants is between 18 and 57 years (average age 25.4 years). The survey was completed by both women (55%) and men (45%). The majority of the participants is well-educated. The share of participants holding a university degree is 51%. Another 42% of the participants achieved degrees with the matriculation standard.

4 Results

In the following section, we present the overview of possible data privacy measures for companies in section 4.1 that resulted from the research process previously described. This overview forms the basis of the presentation of the survey results in section 4.2, that is, the perceptions that customers have of the identified privacy measures.

4.1 Data Privacy Concerns and Measures

The overview of possible data privacy measures is compiled from the literature, legislative texts, and expert interviews. Two publications contain various starting points for measures that can be taken to ease customers' concerns: Morey et al. [33], who describe the role of transparency regarding data collection and usage, and Payne et al. [34], who focus on a list of different laws, regulations, and frameworks, and attempt to reconcile the conflicting agendas of companies and customers. Practical recommendations from Audatis Consulting [35] were used to complement the statements from a practitioner-oriented perspective. Furthermore, we use legislative texts: the European General Data Protection Regulation, which will become applicable law for countries in the European Union in May 2018, the German Bundesdatenschutzgesetz, and the German Telemediengesetz, both finding predominant application with respect to data privacy. To check the completeness of and to verify the previously found statements, we performed three expert interviews in the way described in section 3.1. In the first interview, we talked to an in-house data privacy officer of a German automotive company in order to gain an overview of potential and existing data privacy measures as well as the challenges and difficulties entailed. To verify existing statements and to check whether we had covered all relevant aspects, we conducted a second interview with a researcher who was working on a project with the goal of developing a long-term data privacy strategy for a German bank. To complement our research with input from a legal perspective, we interviewed a lawyer.

From all sources, we collected 141 statements merged to 32 groups. From these groups we derived a particular data privacy measure. All 32 measures can be mapped to one of seven privacy concerns following Smith et al. [15], and as listed in Table **2**.

Table 2. Data privacy concerns presented by Smith et al. [15]

Concern	*Description*
Data Collection	Concern that companies store large amounts of personal customer data.
Data Combination	Concern that customer data from different databases may be combined to gain additional information about a customer.
Internal Secondary Usage	Concern that companies use customer data for a secondary unauthorized purpose.
External Secondary Usage	Concern that customer data are disclosed to a third party and used for a secondary unauthorized purpose.
Errors	Concern that customer data may contain deliberate or accidental errors.
Improper Access	Concern that unauthorized persons are able to view and edit customer data.
Reduced Judgment	Concern that decisions are made in an automated manner and that human intervention in decision-making processes is not possible.

The measures are presented in Tables 3 to 9, grouped by the seven concerns. First, Table 3 represents measures that address customers' privacy concern of Data Collection, meaning that companies might store large amounts of personal customer data.

Table 3. Measures addressing customers' privacy concern of Data Collection

#	*Measure description*
A1	The purpose, scope, and storage time of the data collection and the involved advantages, risks, resulting rights, and obligations are clearly explained to the customer.
A2	Customer data are, as best as is possible, stored anonymously to prevent backtracking of individual customers.
A3	Only the customer data absolutely necessary to provide the agreed service are collected.
A4	Altering or exiting the contractual agreement with regard to personal data is as easy as entering into it. Among others, processing requests occurs quickly and is free of charge.
A5	At the request of the customer and without a long delay, the company provides a set of his personal data free of charge in an easily readable form. Furthermore, the customer has the right to pass these data to other companies.

Table 4 comprises measures that address customers' privacy concern of Data Combination. That is, customer data out of different databases might be combined to gain additional information about a customer.

Table 4. Measures addressing customers' privacy concern of Data Combination

#	*Measure description*
B1	The customer is informed if the company combines his data from various internal and external sources.
B2	If the company combines customer data from various internal and external sources, combination and storage are carried out using anonymous data to prevent backtracking of individual customers.
B3	If customer data are collected for different purposes, the data sets are stored in different databases and are not combined.
B4	The customer decides on whether the company is allowed to combine data from various internal and external sources and can change his decision at any time.

Customers might be concerned that companies use customer data for a secondary unauthorized purpose within the company. Measures addressing the concern of Internal Secondary Usage are listed in Table 5.

Table 5. Measures addressing customers' privacy concern of Internal Secondary Usage

#	*Measure description*
C1	The customer is informed whether and what data are passed on within the company or group of companies and for what purpose.
C2	Customer data are deleted as soon as the original reason for the collection no longer applies or the customer withdraws his permission.
C3	Entering, viewing, altering, and deleting customer data are recorded to make it possible to retrace who changed the data when, and in what manner at any time. The customer can either directly view the log file or is informed about any alterations of his personal data.
C4	If customer data are collected for different purposes, the data sets are stored in different databases and are not combined.
C5	Customers have the opportunity to easily decide which of their personal data are shared with other departments of the company and/or used for other purposes.

Measures addressing customers' privacy concern of External Secondary Usage are presented in Table 6. Customer data might be disclosed to a third party and used for a secondary unauthorized purpose.

Table 6. Measures addressing customers' privacy concern of External Secondary Usage

#	*Measure description*
D1	If customer data are passed on to external third parties, the customer is informed. If customer data are passed on to external third parties, the company ensures that the data are only used in the manner agreed on with the customer through contracts or binding commitments to data protection regulations.

Table 7 cont'd. Measures addressing customers' privacy concern of External Secondary Usage

#	*Measure description*
D3	If customer data are passed on to external third parties, the company or an independent certification organization regularly checks the external third party's compliance with data privacy regulations.
D4	If customer data are passed on to external third parties, data are only forwarded in aggregated or codified form (e.g., income class instead of exact yearly income).
D5	If customer data are passed on to external third parties, the data are – as best as possible – forwarded anonymously.
D6	The company does not pass on customer data to external third parties.
D7	The customer has the choice to easily deny sharing his data with external parties even if doing so results in compromising or the complete abortion of the value delivery.

Customer data might contain deliberate or accidental errors. Measures addressing the concern of Errors are listed in Table 8.

Table 8. Measures addressing customers' privacy concern of Errors

#	*Measure description*
E1	Customer data are checked regularly by the company for completeness, accuracy, and being up-to-date.
E2	The company ensures that no customer data are destroyed or lost by technical and organizational means.
E3	Employees with access to customer data are selected carefully, their behavior is checked regularly, and they are held responsible for malpractice.
E4	Entering, viewing, altering, and deleting customer data are recorded to enable retracing at any time who changed the data when, and in what manner. The customer can either view the log file directly or is informed about any alterations to his personal data.
E5	The customer has access to his data to correct errors, make alterations, or delete data. If he is not provided with direct access to edit his data, they are changed by the company on request.

Table 8 contains measures addressing the concern of Improper Access, which means that unauthorized people might be able to view and edit customer data.

Table 9. Measures addressing customers' privacy concern of Improper Access

#	*Measure description*
F1	If the protection of customer data was violated and their security is at risk, the company immediately informs the customer and the authorities.
F2	Storage and transmission of customer data are protected by technical (e.g., password protection, encryption) and organizational means (e.g., access control, companywide standards regarding handling customer data).
F3	The company ensures that customer data are stored and processed only on its own servers within the European Union or countries trusted by the European Commission.

Customers might be concerned that decisions are taken in an automated manner and that people cannot intervene in decision-making processes, if necessary. This concern, Reduced Judgment, can be addressed by the measures listed in Table 10.

Table 10. Measures addressing customers' privacy concern of Reduced Judgment

#	*Measure description*
G1	The customer is informed whether a decision was made through an automated systems or through an employee of the company. At the customer's request, the reasons for the decision are communicated and explained.
G2	Automated decision processes are continuously tested and checked for deviations.
G3	Decisions that entail legal consequences (e.g., granting a credit) are never made only on the basis of automated systems.

In summary, Tables 3 to 9 represent a comprehensive list of actions that can be taken by companies to mitigate the risk of displeasing customers and to create the potential for delighting customers regarding data privacy.

4.2 Customers' Evaluation of Data Privacy Measures

Companies need to be aware of customers' evaluation of these data privacy measures, which forms the basis for deriving implications for companies' data privacy policies. To determine whether customers consider the implementation of the different identified data privacy measures as "must-be" (basic need), "one-dimensional" (performance need), "attractive" (delighters), or "indifferent," we analyzed the survey's results using the Kano model as described in the previous section. These results are illustrated in Figure 3. Thereby, the measures are numbered as defined in Tables 3 to 9. The abscissa denotes the majority-share of survey participants that determined the measure's classification as one of the four Kano model factors. The ordinate states the spread between the majority-share and the second highest share to evaluate the result's clarity. In the illustration, a green square represents a measure considered to be a basic need by the majority of the participants. Analogously, red dots symbolize measures that are considered to be delighters and gray triangles mark the respective measures as being of indifferent quality. There are no measures considered to be performance needs. To illustrate this approach, we use measure D5 as an example. According to their choice of answers, the majority of the participants (57%, abscissa) see it as a basic need, whereas the second largest group (23%) consider it as a performance need. Thus, the ordinate is 34% (57%–23%), representing a relatively clear result. Overall, the unity among survey participants regarding the classification of a data privacy measure is the smallest bottom left and increases along the bisector. Thus, the distinctiveness of a categorization is highest toward the top right. Valid implications can be derived from the results starting from a spread of at least 10% on the ordinate in Figure **3**.

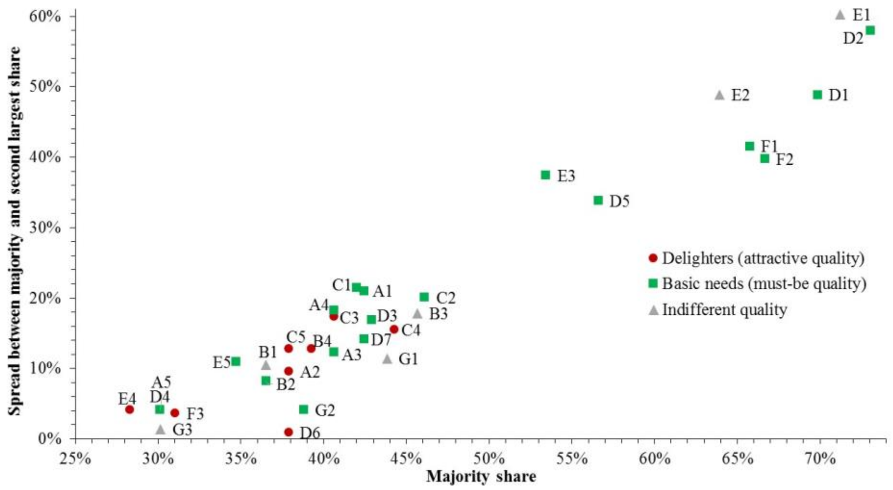


Figure 3. Visualization of the empirical results

Survey participants see 18 out of 32 measures as basic needs. That is, the realization of these measures is neither rewarded nor explicitly demanded by customers. Instead, it is a basic prerequisite when engaging in business with the company. In particular, basic needs can be found among measures addressing the concerns Collection (4 measures out of 5 categorized as a basic need), External Secondary Usage (6/7), and Improper Access (2/3). Hence, these basic needs can be considered a necessary evil because they have downside risk if not implemented but offer no upside opportunities if implemented. The most distinctive example is measure D2, stating that external secondary usage is to be regulated by contracts or other provisions to ensure that data are only used in the manner agreed on with the customer.

Furthermore, no measures are considered to be performance needs. Their constituting properties are, in addition to having a negative impact if not implemented, that they also have the ability to increase customer satisfaction when implemented properly. The total lack of such factors with upside potential is another emphasis of the necessary-evil quality of most data privacy measures.

Entirely, six measures are considered by the survey's participants to be of indifferent quality and, in particular, can be found when addressing the concerns Combining Data (2 measures out of 4 categorized as indifferent) and Reduced Judgment (2/3). These measures do not allow distinctive interpretations toward any direction.

However, there are eight measures categorized as delighters, which are measures that are not required by the customer but may please them, and have no negative impact if not implemented. These measures go beyond the data privacy measures that customers expect. Their implementation positions a company at a level of data privacy commitment higher than anticipated, which has the potential to be rewarded with higher customer satisfaction. Thus, delighters enable companies to differentiate themselves from competitors and to gain a competitive advantage. For instance, Internal Secondary Usage is the concern with the highest share of measures classified as delighters (3/5). In particular, customers can be delighted by providing them with the ability to retrace who changed the data, how, and when (measure C3), or by storing customer data in different and not combined databases, if the data are collected for different purposes (measure C4).

In summary, most of the identified data privacy measures are classified as basic needs. However, survey participants' answers lead to the classification of some data privacy measures as delighters. Thus, our results show that the implementation of data privacy measures has the potential to delight customers.

5 Summary, Limitations, and Future Research

This paper provides an overview of data privacy measures collected from scientific and practitioner-oriented literature, legislative texts, and expert interviews, and can be useful for researchers and practitioners. On top of this overview, this paper provides first insights into customers' perceptions of the identified data privacy measures. By using the Kano model to design a survey with more than 200 participants, we could show that the majority of data privacy measures must be considered as necessary evils for companies. Nevertheless, some data privacy measures can even delight customers. Thus, this paper's result is that certain data privacy measures have the potential to increase customer satisfaction and enable a competitive advantage for companies. Accordingly, researchers and practitioners may use our approach as inspiration when deriving a data privacy strategy because evaluating customers' perception may assist in prioritizing the implementation of data privacy measures. Measures classified as basic needs should be implemented by every company to avoid data privacy incidents and negative effects on customer satisfaction. Companies that strive for delighting customers through data privacy may also implement measures classified as delighters.

However, researchers and practitioners need to be aware of our research having some limitations. First, the research approach is limited to the consideration of a specific aviation sector scenario. To verify the general validity of the conclusions, the survey has to be rerun for further settings that refer to other industries. Second, in the field of data privacy, statements of customers in empirical surveys do not necessarily match their actions in the real world. According to Norberg et al. [36] and Acquisti and Grossklags [37], the so-called privacy paradox describes the discrepancy between customers' intentions to protect their own privacy and their real-world behavior. To take into account this phenomenon, the results of the survey should be verified in real-world situations. Third, in general, the classification of delighters is less clear than the classification of basic needs. That is, when interpreting this paper's results, implications must be challenged according to the principle of prudence. When in doubt, a measure should rather be considered a basic need than being of indifferent quality or a delighter. Future research could follow Matzler et al. [18], who state that unclear results spread out over several categories can be a starting point for market segmentation. Thus, further research could examine the categorization of data privacy measures as Kano model factors depending on demographic characteristics.

When providing an overview of data privacy measures and outlining the potential to increase customer satisfaction by applying certain data privacy measures, we could also point out main areas of further research relevant to both researchers and practitioners. Specifically, we plan to extend our research to other industries to evaluate general validity in the near future. Further research can also focus on a break-down of single data privacy measures into its individual components and the influence of these granular aspects on customers' satisfaction with a particular data privacy measure.

References

1. Matthing, J., Sandén, B., Edvardsson, B.: New service development: learning from and with customers. International Journal of Service Industry Management. 15, 479–498 (2004)
2. Saarijärvi, H., Grönroos, C., Kuusela, H.: Reverse use of customer data: implications for service-based business models. Journal of Service Marketing. 28, 529–537 (2014)
3. Berendt, B., Günther, O., Spiekermann, S.: Privacy in e-commerce: stated preferences vs. actual behavior. Communications of the ACM. 48, 101–106 (2005)
4. Preibusch, S., Kübler, D., Beresford, A.R.: Price versus privacy: an experiment into the competitive advantage of collecting less personal information. Electronic Commerce Research. 13, 423–455 (2013)
5. Acquisti, A., Friedman, A., Telang, R.: Is There a Cost to Privacy Breaches? An Event Study. In: ICIS 2006 Proceedings, pp. 1563–1580 (2006)
6. Muntermann, J., Roßnagel, H.: On the Effectiveness of Privacy Breach Disclosure Legislation in Europe: Empirical Evidence from the US Stock Market. In: Jøsang, A., Maseng, T., Knapskog, S.J. (eds.) NordSec 2009. LNCS, vol. 5838. pp. 1–14. Springer, Heidelberg (2009)
7. Tanner, A.: Here Are Some Of America's Most Privacy Friendly Companies. http://www.forbes.com/sites/adamtanner/2013/09/11/here-are-some-of-americas-most-privacy-friendly-companies/#1684385b306a (Accessed: 02.11.2016)
8. Buhl, H.U.: IT as curse and blessing. Business & Information Systems Engineering. 5, 377–381 (2013)
9. Ahmad, A., Mykytyn, P.: Perceived Privacy Breach – the Construct, the Scale, and its Antecedents. In: AMCIS 2012 Proceedings. pp. 1–9. (2012)
10. Pavlou, P.A.: State of the information privacy literature: where are we now and where should we go? MIS Quarterly. 35, 977–988 (2011)
11. Nicholas-Donald, A., Matus, J.F., Ryu, S., Mahmood, A.M.: The Economic Effect of Privacy Breach Announcements on Stocks: A Comprehensive Empirical Investigation. In: AMCIS 2011 Proceedings. pp. 1–15. (2011)
12. Cavusoglu, H., Mishra, B., Raghunathan, S.: The effect of Internet security breach announcements on market value: capital market reactions for breached firms and Internet security developers. International Journal of Electronic Commerce. 9, 69–104 (2004)
13. Campbell, K., Gordon, L.A., Loeb, M.P., Zhou, L.: The economic cost of publicly announced information security breaches: empirical evidence from the stock market. Journal of Computer Security. 11, 431–448 (2003)
14. Hovay, A., D'Arcy, J.: The impact of denial-of-service attack announcements on the market value of firms. Risk Management and Insurance Review. 6, 97–121 (2003)
15. Smith, H.J., Milberg, S.J., Burke, S.J.: Information privacy: measuring individuals' concerns about organizational practices. MIS Quarterly. 20, 167–196 (1996)
16. Klingspor, V.: Why Do We Need Data Privacy? In: Michaelis, S., Piatkowski, N., Stolpe, M. (eds.) Solving Large Scale Learning Tasks. Challenges and Algorithms. LNCS, vol. 9580. pp. 85–95. Springer (2016)

17. Buchmann, E., Böhm, K., Raabe, O.: Privacy 2.0: Towards Collaborative Data-Privacy Protection. In: IFIP International Conference on Trust Management, pp. 247–262. Springer US (2008)
18. Sarathy, R., Robertson, C.J.: Strategic and ethical considerations in managing digital privacy. Journal of Business Ethics. 46, 111–126 (2003)
19. Nufer, G., Prell, K.: Operationalisierung von Kundenzufriedenheit. Reutlinger Diskussionsbeiträge zu Marketing & Management. 4, 1–18 (2011)
20. Parasuraman, A., Zeithaml, V.A., Berry L.L.: A Conceptual Model of Service Quality and Its Implications for Future Research. Journal of Marketing. 49, 41–50 (1985)
21. Ladhari, R.: A review of twenty years of SERVQUAL research. International Journal of Quality and Service Sciences. 1, 172–198 (2009)
22. Hackl, P., Westlund, A.H.: On structural equation modelling for customer satisfaction measurement. Total Quality Management. 11, 820–825 (2000)
23. Bartikowski, B., Llosa, S.: Customer satisfaction measurement: comparing four methods of attribute categorisations. The Service Industries Journal. 24, 67–82 (2004)
24. Füller, J., Matzler, K.: Customer delight and market segmentation: An application of the three-factor theory of customer satisfaction on life style groups. Tourism Management. 29, 116–126 (2008)
25. Löfgren, M., Witell, L.: Two Decades of Using Kano's Theory of Attractive Quality: A Literature Review. The Quality Management Journal. 15, 59–76 (2008)
26. Lai, H.J., Wu, H.H.: A Case Study of Applying Kano's Model and ANOVA Technique in Evaluating Service Quality. Information Technology Journal. 10, 89–97 (2011)
27. Arbore, A., Busacca, B.: Customer satisfaction and dissatisfaction in retail banking: Exploring the asymmetric impact of attribute performances. Journal of Retailing and Consumer Services. 16, 271–280 (2009)
28. Matzler, K., Hinterhuber, H.H., Bailom, F., Sauerwein, E.: How to delight your customers. Journal of Product & Brand Management. 5, 6–18 (1996)
29. Kano, N., Seraku, N., Takahashi, F.: Attractive quality and must-be quality. The Journal of the Japanese Society for Quality Control. 14, 147–156 (1984)
30. Mikulić, J., Prebežac, D.: A critical review of techniques for classifying quality attributes in the Kano model. Managing Service Quality: An International Journal. 21, 46–66 (2011)
31. Strategy&, http://www.strategyand.pwc.com/perspectives/2015-aviation-trends (Accessed: 18.08.2016)
32. Harris, E.C.: Personal data privacy tradeoffs and how a Swedish church lady, Austrian public radio employees, and transatlantic air carriers show that Europe does not have the answers. American University International Law Review. 22, 745–799 (2007)
33. Morey, T., Forbath, T., Schoop, A.: Customer data: designing for transparency and trust. Harvard Business Review. 93, 96–105 (2015)
34. Payne, D., Landry, B.J.L., Dean, M.D.: Data Mining and Privacy: An initial attempt at a comprehensive code of conduct for online business. Communications of the Association for Information Systems. 37, 717–732 (2015)
35. Audatis Consulting, https://www.audatis.de/wp-content/uploads/Checkliste_Datenschutz_TOM_nach_9_BDSG.pdf (Accessed: 18.08.2016)
36. Norberg, P.A., Horne, D.R., Horne, D.A.: The privacy paradox: personal information disclosure intentions versus behaviors. Journal of Consumer Affairs. 41, 100–126 (2007)
37. Acquisti, A., Grossklags, J.: Privacy and rationality in individual decision making. IEEE Security & Privacy. 2, 24–30 (2005)

Social Network Services: Competition and Privacy

Claus-Georg Nolte, Jonas Schwarz, Christian Zimmermann

University of Freiburg, Institute of Computer Science and Social Studies,
Department of Telematics, Freiburg, Germany
{nolte,schwarz,zimmermann}@iig.uni-freiburg.de

Abstract. Social Network Services (SNS) business models highly depend on the gathering and analyzation of user data to obtain an advantage in competition for advertising clients. Nevertheless, an extensive collection and analysis of this data poses a threat to users' privacy. Based on an economic perspective it seems rational for Social Network Operators (SNO) to ignore the users' desire for privacy. However, privacy-friendly services might have the potential to earn users' trust, leading to an increased revelation of personal data. Addressing these issues, we examine the existing privacy problem in SNS in the context of competition between SNO to investigate whether competition tend to enhance user privacy or whether it is the root of its violation. Therefore, this paper investigates the interconnectedness of the market structure and privacy problems in SNS. After analyzing the users' and the advertisers' side of SNS, their competitiveness and its influence on user privacy are examined.

Keywords: Privacy, Social Network Services, Competition, Multi-Sided Platforms, Two-Sided Markets

1 Introduction

At least since Facebook was published in 2004, social network services (SNS) have constantly been on the rise and consume evermore of our daily online time and, thereby, of our personal data. Moreover, since the beginning of the smartphone age SNS have even been following us from our desktops to every place we go. They demand us to share every bit of our lives with our friends within the network and, thus, with the network itself. Consequently, Harvard Law professor Jonathan Zittrain deduced in 2008 that this technology threatens "to push everyone towards treating each public encounter as if it were a press conference" [1]. This thirst for user data can be explained with the business model of SNS which heavily depends on gathering and analyzing user data to deliver targeted advertisements [2]. This extensive collection and analysis of user data poses a severe threat to users' privacy. Hence, Margo Seltzer, professor in computer systems at Harvard University, sets it straight at the Davos Forum in 2015 stating that "privacy as we knew it in the past is no longer feasible" [3].

Disclosed data itself is necessary for SNS businesses to improve their targeting of advertisements and thereby obtain an advantage in competition for business customers,

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Nolte, C.-G.; Schwarz, J.; Zimmermann, C. (2017): Social Network Services: Competition and Privacy,
in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung
Wirtschaftsinformatik (WI 2017), St. Gallen, S. 822-836

which demand for precisely targeted advertising. Hence, users' demand for privacy-friendliness only becomes important insofar as it is helps providers to gain trust and to get users to reveal even more personal data [4]. Thus, it seems not an irrational ignorance of users' desire for privacy by the providers but a rational choice in an economic competition not to prioritize this desire.

The paper takes an economic perspective on the present privacy problem in SNS to investigate whether competition between providers tends to enhance user privacy or whether it is the root of its violation. For this purpose we investigate privacy in those business focusing on the market structure, thereby taking into account that SNS constitute multi-sided platforms (MSP) [5]. Analyzing the users' and the advertisers' side of SNS, we compare their competitiveness and its influence on user privacy. Therefore, we build upon insights from theoretical literature concerning MSP, behavioral research papers about SNS and privacy, as well as market evidence.

In the following we will first state the essential definitions for our investigation in section 1.1 and further give a brief overview of the related literature in section 2. Thereafter, we start with the analysis of the influence of SNS competition on user privacy, firstly by examining the characteristics of the goods which are up for rivalry on the different market sides in the SNS environment, and secondly, by investigating those goods and competitions in detail considering appropriate findings from scientific literature (see section 3.1). Thirdly, we will analyze their impact on user privacy in section 3.2 and following. Finally, we will discuss our results, match them with empirical evidence (see section 4) and give a summary of our paper in the conclusion.

1.1 Essential Definitions

As mentioned above, we aim at investigating the interrelation of the market structure in social networks services (SNS) and user privacy. Hence, there is a need to clarify the term and interpretation of SNS. According to the updated definition of Kane et al. [6], which builds upon the earlier characterization of Ellison [7], SNS contain the following features: "users (1) have a unique user profile that is constructed by the users, by members of their network, and by the platform; (2) access digital content through, and protect it from, various search mechanisms provided by the platform; (3) can articulate a list of other users with whom they share a relational connection; and (4) view and traverse their connections and those made by others on the platform" [6]. In the following we use this definition as a basis to describe and understand SNS. However, we add the restraint that the main revenue source of the SNS should be advertising to ensure the multi-sided platform character of our investigation object. Hence, the scope of our analysis contains SNS such as Facebook, Google+ and Twitter.

Moreover, we define the term of Social Network Operators (SNO) as companies, which "provide the underlying basic services and infrastructures, needed by users to interact with each other" [8]. In the case of Facebook, the website facebook.com constitutes the SNS while the company Facebook Inc. is the SNO.

Further, the term multi-sided platform, also known as two-sided platform or market, requires a clear definition, too. Therefore, we draw on the work of Staykova and Damsgaard [5]. According to their research, MSP (1) enable direct interaction between two or more participants affiliated to them; (2) are containing homing and switching costs for those participants; and (3) include direct and indirect network effects [5] (see Fig. 1). In our setting, homing costs are the costs in money, effort, time and other aspects of entering and using an MSP. Moreover, switching costs are similar onetime costs that occur if participants switch from one platform to another. Additionally, network effects occur when the value of the platform or its product for one participant is influenced by the total number of participants (of the same or another side). Those effects can be both positive or negative and are seen as a key aspect of MSP [9].

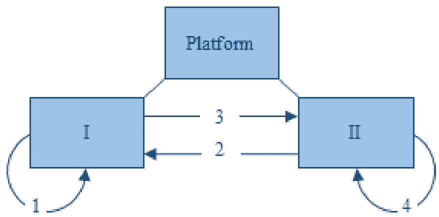


Figure 1. MSP model, including platform participants (I&II), same-side network effects (1&4), and cross-side network effects (2&3) [5].

The privacy definition used in this paper builds upon Westin's definition of privacy. Hereby, he defines it as "the claim of individuals, groups or institutions to determine for themselves when, how, and to what extent information about them is communicated to others” [10]. For him, “this, also, involves when such information will be obtained and what uses will be made of it by others” [11]. In the context of SNS, we consider privacy as the capability of SNS users to control their personal data and its collection, aggregation, analysis and possible transfer by the SNO and third parties, as well as users' ability to optimize the amount of data disclosure and its security against misuse with respect to their preferences.

The legal definition for personal data is provided by the EU data protection directive as “any information relating to an identified or identifiable natural person”.[1] However, for our case a broader classification is useful. Drawing from Nolte, we define personal data as “any data revealed by user action, starting from simple likes, to direct personal information, and even analysis of users click and browse behavior” [12]. This diversification is necessary to capture that by the aggregation and combination of meta- and behavioral data, inferences can be drawn. Those can constitute sensitive information for users and sometimes even personal data in the definition of European data protection law [13]. Given this definition, it is obvious that the data-centric business model described above constitutes a severe threat to users' privacy.

[1] EU Data Protection Directive (95/46/EC), 1995.

2 Related Literature

User privacy in SNS is a widely examined and discussed issue. Basically, the topic can be divided into two main research streams: Behavioral and user-focused research as well as provider-focused research. The former offers the seemingly contradictive result that users on the one hand care for privacy and try to preserve it with privacy-seeking behavior [14], but on the other hand do not act privacy-aware and seemingly carelessly disclose personal data when using Internet services [15]. Those findings led to the definition of the so-called "privacy paradox" [16]. Furthermore, provider-focused research has shown that privacy is not a major market factor in the competition for user attraction although users estimate it of high importance for them [17, 18].

However, the question wherefrom this dynamism arises which drives SNS providers to claim more and more user data and thereby restrict user privacy is still open to research. Different forces interact which each other and the participants in an SNS. Most of those forces have direct or indirect influences on privacy [19]. One popular assumption is that users' demand for privacy is of minor priority for SNS providers because users are not willing to pay for it [20] and the monetary income is generated by advertisement customers [2]. However, recent successful mail services show that a minority of users is willing to spend small monetary amounts for increased communication privacy.[2] Other companies in the Internet search business even display the possibility to succeed without demanding any money for a privacy respecting service [21]. Nevertheless, this willingness to pay for privacy either in a monetary way or in terms of switching costs seems of no significance for SNS [15, 22, 23].

Further, the more general question of competition in MSP has been addressed from different angles. The economics of two-sided markets have most notably been explored by Rochet and Tirole [24]. Further research into MSP market structures has been conducted by Armstrong [25], Evans and Schmalensee [26] and others [5, 27]. While these highly recognized works provide deep insights into the economics of MSP, they do not cover privacy issues. A variety of publications address the questions of competition and monopolistic tendencies in online MSP, while focusing on the search engine market and Google's market position in particular [28, 29] or on SNS and Internet services in general [30]. However, the potential interrelation of the privacy problems in SNS and the market structure are not in the focus of current research.

3 Economic Analysis

In the following section we introduce and clarify the MSP business of SNS. Further, the traded goods in an SNS environment are examined from different angles to determine their competitive character. Afterwards, we consider their influences on user privacy for both sides of the SNS/MSP entity, namely the users' and the advertisers' side. For simplicity other SNS participants, like application developers, are bypassed.

[2] Mailbox.org, Posteo.de and others.

As argued above, SNS in general constitute an MSP, generating revenue by brokerage of targeted advertising to its users for business partners [2, 31, 32]. A closer view on the market structure reveals strong direct same-side network effects between its users, because each additional user makes the SNS more attractive to others [33, 34]. Moreover, there are indirect cross-side network effects between users and advertisers. Each additional user makes the network more valuable for advertising clients. This is due, firstly, to a broader audience for targeted advertisement and, secondly, to a higher amount of user data, which elicits the possibility of drawing inferences and thereby creates more precise profiles [5]. On the other side, users are at least accepting personalized advertisements as a price to use SNS free of monetary charge [32]. However, there are no positive network effects between the advertisers, rather the opposite can be assumed. While advertisers profit from additional SNS users and often perform side-advertisement by promoting their SNS company profiles (e.g. advertising the company's Facebook-page or Twitter-account), they are rivals to other advertisers within the same network for the limited space of targeted advertisements (see Fig. 2).

3.1 Features of Social Network Service Goods

As already stated, advertising companies are in rivalry for the limited space for targeted advertisements. More precisely, currently leading SNS auction targeted advertisement for specific audiences or keywords in a real-time bidding system between interested advertisers. The space for advertisement in the network is limited and advertising clients can exclude each other through a higher bid for the same keyword or target group, thus, the good of advertisement is exclusive and a rival good. Furthermore, a SNO can decide to exclude some advertisers from its service. Hence, advertisement in SNS is a classic private good (see Table 1). Thus, one has to assume that there exists a strong rivalry between similar advertisers. Classifying this insights into a feature of goods table shows that the service of providing targeted advertisement within SNS is to be considered a private good [35].

Categorizing the SNS users-side is more complicated. First of all, we have to distinguish between two different goods: the plain SNS membership and the actual usage of the network. The first requires usually only a valid email address and roughly two minutes for filling out the application form and confirming one's mail address. The second comprises the aforementioned homing costs, time and effort to understand the SNS' practice as well as adding user created content. Both add up to the aforementioned switching costs (c.f. section 1.1). We already argued that users experience same-side network effects from other users. Hence, profile creating and SNS usage is non-rival. Creating a profile and actively participating in an SNS prevents no one else from joining or using it.[3] In addition, the question to be answered is whether those user-sided goods are excludable or not which makes the difference between a public good and a club good. At first glance, it seems intuitively to argue that those goods are public because no one seems to be able to exclude someone else from the usage. However, people from Turkey trying to access Facebook during the military coup in July 2016 or generally

[3] Except for server overload which is not discussed here for simplicity.

trying to use it in China or North Korea will disagree. Countries as well as SNO have technical instruments to restrict SNS access or directly ban specific users.[4] Hence, the SNS registration and usage is non-rival but excludable for users' side whereby it complies with the characteristics of a club good (see Table 1).

Table 1. Feature of Goods Classification for SNS

	excludable	**non-excludable**
rival	*"private good"* targeted advertisement	*"common good"*
non-rival	*"club good"* SNS registration & usage	*"public good"*

3.2 The Social Network Operator Viewpoint

The last angle missing is the SNO perspective. First, the providers compete among themselves for advertising clients. All SNS are offering roughly the same product on this market side: targeted advertisement. It applies here that neither can money spent by an advertiser in one network be spent twice, nor can advertisement space be assigned multiple times. Accordingly, one can assume that the SNS advertisement market side is in strong competition because several providers supply a comparable private good to a high quantity of advertisement-willing companies (see Fig. 2).

Second, user registration and membership seems non-rival from an SNO perspective because users can easily set up multiple accounts in different SNS. Moreover, a provider is not able to prohibit its members from registering at other networks nor to hinder other services to open their registration for them. Yet, users can decide to refuse a certain SNS. However, competition undoubtedly exists for attracting users between SNO, since the quantity of accounts is a signal to attract advertisers.

Third and most interesting from an SNO perspective is the time users spend in the network. It seems to be a highly valuable good for providers, because it increases the possible quantity of advertisements shown to users and probably the amount of data disclosed by them [14]. Further, users' time and attention is limited and can only be spent once to an SNS. A strong competition for users' time between SNS and also other services can thus be presumed [36]. As a result, users are not only paying for an SNS with revealed personal data but also with their time and attention (see Fig. 2).

[4] Technical workarounds for users are neglectable for our analysis.

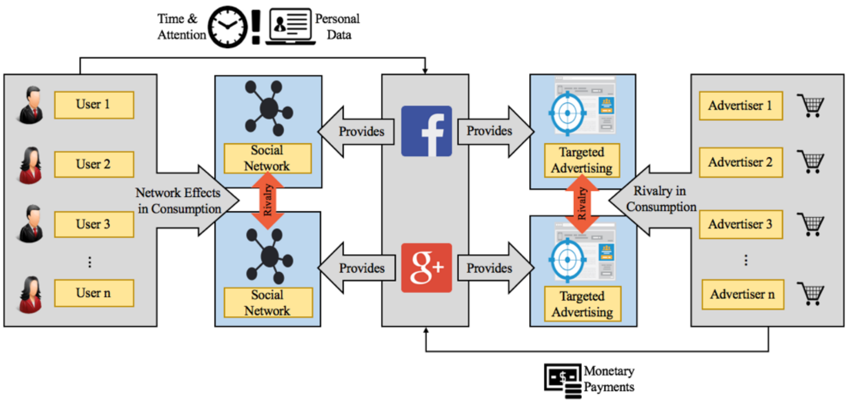


Figure 2. The market structure of SNS

3.3 Competition on Users' Side

Analyzing the SNS' competition on the users' side of the market structure, we firstly identified two relevant factors: trust and enjoyment [33, 34]. On the one side, as described before, the SNO wants to receive users' time and therewith their attention for advertisements and their disclosed personal data for improved targeting of those advertisements [37]. On the other side, users want to enjoy an SNS and demand it trustworthy [4, 33, 34], while enjoyment also includes strong same-side network effects, i.e. finding the own friends within the same SNS [34].

Targeting the trust factor first, literature shows that increasing trust in their SNS can be achieved for SNO by implementing privacy controls [33, 38]. This is not only relevant because improved trust increases the SNS usage but also the quantity of disclosed data and the acceptance of advertising [39, 40]. Hence, trustworthiness is beneficial for SNS to bind users to the service as well as to receive more user-generated content and reliable user data. Undoubtedly the implementation of privacy controls has a positive influence on user privacy in SNS [19]. However, findings suggest a design conflict between privacy and usability and, thus, enjoyment [18].

Talking about the enjoyment of an SNS, it seems to be the most crucial factor in SNO competition. Firstly, because it attracts users to join a network and, thus, self-evidently increases positive same-side network effects between the users as well as positive cross-side network effects from the users' side to the advertisers'. Secondly, enjoyment tempts users to spend more time with the network, leading again to positive influence on advertisers. Additionally, more time spent in an SNS also increases the quantity of user-generated content because content creators value a platform more if they have a larger audience [41]. Further, a higher amount of content again attracts more advertisers because it enhances the providers' targeting ability for advertisement. To conclude, we find plenty of motives for providers to compete with enjoyment for user registrations and user time. SNO are facing this competition by increasing their own platform stickiness. Literature results show that there exist mainly two ways of doing so which add to another: increasing the content of the platform and implementing more features and functionalities to it [5, 36]. While the latter seems at first glance privacy neutral for users, changing the platform appearance to entice users to enter more data clearly is a threat to their privacy [14]. However, additional platform features also contain the potential to harm user privacy when they educe additional personal data from users or even leak those data from the platform to third parties if the SNO decides to open her network to external application developers.

3.4 Competition on Advertisers' Side

As already stated, SNO compete for selling targeted advertising to corresponding customers (c.f. section 3.2). To attract those advertisers, we identified four factors as most relevant: the quantity of users, the accuracy in user targeting to serve the advertisements, the time users spend in the network and the price to advertise to the targeted user group. In their evolution, the majority of SNS followed the same path: first starting one-sided and attracting users, and after hitting a critical mass, implementing advertisement and thereby evolving into a two-sided platform [5]. Later, most SNS also included external application developers and other services (e.g. identity management) and, thus, transformed into an MSP. However, we already covered the privacy impacts of competition for user registration in the previous section.

Besides the pure quantity of users, the average time a user spends in the network appears to be the crucial factor in the competition for advertisers [36]. Its theoretical competition and privacy impacts on users' side were already shown above. However, another option for tying users closer and longer to a network seems to enhance the SNS content [6, 42]. This can either happen by tempting users to post more data or by including external content creators and their content directly into the network (e.g. news sites or celebrities). A further method is simply to acquire competitors and include their services into the own SNS. We already showed the possible privacy threats of tempting users to reveal more data above. The method of including further content form external content creators is basically privacy neutral, except for users' active reaction on this content (e.g. likes and comments). However, implementing bought-up competitors and especially merging their existing user data and accounts with already existing in-network user data can be extremely privacy invasive. Merging that data and drawing inferences from the new database can reveal information which the user initially wanted to hide by audience-segmentation via using two separated services.

The third identified factor is the accuracy of user targeting. The most obvious way of improving this factor is to gather more user data, either directly from the users or from external sources, to analyze it with algorithms. The impacts of user data have already been discussed. However, one has to expect that the possibilities of enhancing targeting by user data somewhere hits a limit value where gathering more data does not result in any more improvements. Hence, another method to improve the targeting constitutes a direct inquiry of users either for their interest or indirectly by giving them the controls to correct their information and drawn inferences connected to their account [37]. The latter can be privacy enhancing for users if it is allowed for them to delete data or at least exclude certain information from the targeting mechanisms.

Finally, the cost-benefit-factor of advertising in an SNS depicts a crucial aspect, since the advertisement side is the monetary paying side. As every agent acting in an economic environment, advertisers seek for the most efficient way to spend their money and, thus, to target their audience. Besides the developers' achievement of a leading targeting algorithm and the already treated topic of feeding it with user data, SNO have another factor to influence the efficiency of their advertisement offer: economics of scale. Additionally to non-rivalry and excludability, user membership and activity in SNS have the characteristic of low marginal costs. The costs to provide the service to an additional user are, after establishing a working system, marginal for the SNO. The same holds true for the advertisers' side and the service of automated advertising space auctions. Hence, this leads to rising returns of scale which makes an SNS the more efficient the larger both sides are. Thus, the direct way to increase SNS attractiveness for advertisers is to gain more users joining and spending time in the SNS. This becomes even more efficient if those users originate from different target groups, because the SNS then represents all sections of society. This insight leads us directly back to section 3.3 and the discussed influences on privacy and competition.

3.5 The Trump Side of Competition

In the previous sections, we analyzed the two major sides of SNS regarding competition and their impacts on user privacy. Table 2 summarizes the different activities of SNO and their influences. However, it seems uncertain which side in competition outweighs the other and, thus, whose needs are favored by the SNO for economic reasons. If the users' side and therewith users' needs are preferred it is to be expected that the trust factor could lead to an improvement of users' data control options and, thus, to an enhanced user privacy. We expect the opposite if advertisers are the SNO-favored SNS side, due to the demand of evermore user data for better profiling.

Table 2. SNO activities and their influence on user privacy

SNO Activity	Influence on User Privacy
Implementing Privacy Controls	+
Give Users Control to Correct and Enhance the Information and Drawn Inferences Connected to their Account	+
Implement more Features and Functionalities	❍
Changing Platform Design to Entice User to Enter More Data	-
Merging their Existing User Data and Accounts with Already Existing In-Network User Data	-

Following the influential papers on MSP, the standard link from classical economics between the inverse relation of price over marginal cost and elasticity of demand does not hold true for MSP. In other words, the service on one side will be served by the provider even if this service and the price paid for it by the its participant alone is not profitable [26]. Hence, the loss from one side has to be outweighed by the profits from the other. To identify the provider's cash cow we have to find out on which side multihoming is most prevalent [24, 25]. In other words, the side where the MSP participants use more than one platform simultaneously is expected to be overpriced, while the singlehoming side is expected to be subsidized [43].

4 Discussion

In this section we will match the theoretical analysis of competition and its influence on privacy in SNS with available market observations and empirical evidence. Subsequently, we discuss whether this evidence confirms our theoretical findings.

First, targeting the mentioned trust factor of section 3.3 we find that current developments indicate that SNS and other Internet services recognized the coherency between users' trust and user-generated content. Respectively, Facebook introduced new privacy controls in 2008 [44] and constantly improves them [45] while Google implemented its own user privacy controls lauded by specialized press [46]. However, another reason for this privacy trend in Internet services might be the upcoming reform of the EU Data Protection Directive which comes with rigid laws and harmful financial penalties in case of violations [47]. Hence, the future of this trend is sustainably influenced not only by user behavior but also by the corresponding law. Utterances of the European Commissioner let us assume that the end of the road of regulating Internet services concerning competition and user privacy hast not been yet reached [48].

As showed in sections 3.3 and 3.4, SNO compete against each other for users' time spent on their platform as well as for content. Therefore, they implement new features into their platforms or integrate taken-over services and external apps. Recent developments in Internet services and SNS make this competition visible. The acquisition of WhatsApp by Facebook, as well as the takeover of Instagram led to a domination of Facebook in the branches of mobile messaging and mobile photo sharing. Moreover, Facebook started partly merging Facebook and Instagram accounts and announced in the latest terms and conditions change of WhatsApp that phone numbers and contacts will be transferred to Facebook [49]. Both clearly contain the aforementioned privacy threats of merging different services accounts.

As a result, Facebook's recent introduction of instant articles can not only be interpreted as an attempt to enlarge the own content but also as an attack on Google and Twitter [50]. Furthermore, Facebook included the feature of selling tickets for events and recently announced the possibility to run crowd-funding campaigns and collect money directly inside the SNS [51]. This seems to be an attempt to use its already large user base and the resulting network-effects to include the markets for online ticket-sale and crowd-funding and, thereby, keep the users as long as possible in the Facebook environment. The same applies to implementation of an own browser within the Facebook app, a strategy also used by Twitter.

Moreover, companies like Google, Facebook and Twitter provide identity management features which enable users to log-in with their already existing SNS accounts to external services. What seems like a comfortable feature to make users' life of managing different online accounts easier, can also be interpreted as a way to gather more data from external services, track users beyond the own platform and enhance both targeting and time for displaying advertisement. We already pointed out the privacy threat aspects of those strategies.

However, the most interesting question is which side, the users' or the advertisers', is subsidized and who is the "cash cow" (see section 3.5). Evidence from the magazine industry suggest that users are subsidized and advertisers are the main income source [52, 53]. The fact that users are not paying any monetary price for SNS and advertising is the main income source for SNO strongly supports this view [54]. Nevertheless, considering multihoming as the crucial factor, recent statistics suggest that it is on the rise on users' side with 52% of US users using two or more social media sites in 2014 [55]. However, the relevant factor of SNS competition on users' side is also the time spent in the network (see section 3.2). Facebook leads by far with 70% of its users using the platform daily before Instagram with 49% and Twitter with 36%, while the second also belongs to the Facebook environment. Moreover, "the engagement of Facebook users continues to grow, while daily use on other platforms shows little change" [55].

Because, we lack numbers for multihoming behavior on the advertisers' side we use the market distribution as an indicator. In 2014 Alphabets' share of the net digital advertising revenue was 31%. This revenue includes the income from targeted advertising on the Google search sites and also the advertisement revenue from services like YouTube, Google+ and AdWords. The leader is followed by Facebook with a market share of nearly 8% and the Chinese online search engine Baidu with close to 5% [56]. Despite these distinct numbers the development of the online advertisement market indicates that the competition between the targeted advertisement-offering SNO is rising. While Alphabet was able to keep its market share over the last three years, competitors are catching up. Facebook nearly doubled its market share from 2012 to 2014 from 4% to close to 8%. Besides, except for two, all other market participants beyond the 0,5% market share were able to claim slightly more percentages each year [56]. This development indicates not only that online targeted advertisement is becoming more popular but also let us presume that adverting clients tend to use more than one Internet services and, thus, show multihoming behavior.

The point that both sides of SNS seem to show multihoming behavior and that the users' side does not pay for their usage in terms of money makes it difficult to apply the insights of MSP markets and multihoming here. As mentioned afore, on the one hand one could claim that users are the subsidized side because they are enjoying SNS for free. However, on the other hand one could also argue that advertisers are being favored and thus subsidized because they are the crucial revenue source of SNO. Moreover, advertisement prices in SNS seem comparably low to those paid in print media and we expect the targeting to be more exact due to the revealed user data [54]. Hence, advertisers could be subsidized with better targeting for comparably lower prices while users could be overpriced in terms of private data elicit by the SNS and, thus, lower privacy. However, according to the lack of literature and data available we are not able to finally assess this case.

In summary, our investigation shows that competition between SNO does neither necessarily improve user privacy nor does it generally harm it. Competition in MSP like SNS is complex and our analysis shows that there are indeed various privacy harmful aspects. However, the competition for users could have privacy friendly consequences, if the trust factor outweighs the privacy contrary implementation of additional features and overtaken services.

4.1 Limitations & Future Research

Our work is limited due to the available data about certain market aspects of SNS competitions. We found no data about multihoming behavior of advertisers regarding the purchase of targeted advertising in SNS. Additionally, there is no evidence whether newly integrated privacy controls by Facebook and Google are actively used and, thus, increase user privacy or if they are just having a trust-building and thereby possible privacy harming effect. Moreover, the comparably low costs of targeted advertisement in SNS can partly be explained by the strong economics of scale and the near-zero variable cost of running an SNS and displaying advertisement compared to print media. Hence, future research in this area should tackle and close this data gap.

Furthermore, considering our classification of SNS goods, club goods as the SNS membership and usage from user side also have the characteristic of low marginal costs. This leads to rising returns of scale, which makes club goods ideal for natural monopolies [35]. Considering the costs to set up an SNS by programming the service and establishing the computing power to serve a broad user base as well as the high costs of running such a server infrastructure as fixed costs, there exist monopoly tendencies for this market [57]. Future research could build upon our work to investigate those tendencies and their influence on privacy for the SNS market.

5 Conclusion

In this paper we investigated the challenge of user privacy in SNS from an economic perspective. The aim was to analyze if competition between networks tends to decrease or enhance user privacy. Therefore, we first clarified that SNS are multi-sided platforms with at least two sides: users on the one and advertising clients on the other side. Furthermore, we characterized the traded goods within SNS markets from the three viewpoints of users, advertisers and social network operators showing that SNS membership and usage are club goods from user perspective, while targeted advertisement is a private good from an advertiser and SNO viewpoint.

Moreover, SNO compete for advertisers on the one side. On the other side they compete for users' membership and more intensely for users' time. In order to attract the advertisers, providers have to maximize their targeting for advertisements and the time users spend within their platform. Hence, they have the incentive to gather as much personal data from users as possible and enhance their platform with additional content, features and services to bind users' attention and spend time in the SNS. All analyzed factors of competition in the SNS environment contain unilateral privacy threatening aspects. Except for the factor of user trust for an SNS, gained by implementing privacy controls as one aspect to get users to reveal more data. Furthermore, the fact that users do not pay in monetary terms for SNS usage but that advertising clients are the crucial revenue source for SNO suggests that users might be discriminated against and overpriced in terms of personal data disclosure. However, the validation by the MSP theory where the multihoming side is overpriced and the singlehoming side is subsidized provides no sufficient results to verify or reject this assumption. In summary, this analysis shows that competition in the SNS environment does neither generally harm user privacy nor does it necessarily improve it. However, the latter seems less likely, unless the competition for users' trust outweighs all the other privacy divergent aspects. The case of competition in the Internet MSP is complex and current statistics provide not enough data to give a solid answer to the question whether advertisers or users are subsidized by the SNO. Nevertheless, the analysis of our research question shows that competition can be assumed to have a negative influence on user privacy at the present stage.

References

1. Zittrain, J.: The future of the internet - and how to stop it. Yale University Press (2008)
2. Beuscart, J.-S., Mellet, K.: Business Models of the web 2.0: Advertising or the Tale of Two Stories. Communications & Strategies, Special Issue (2008)
3. Seltzer, M.: The Surveillance State of Today. Davos, Swiss (2015)
4. Lawani, O., Aïmeur, E., Dalkir, K.: Improving Users' Trust Through Friendly Privacy Policies: An Empirical Study. International Conference on Risks and Security of Internet and Systems, 55–70 (2015)
5. Staykova, K.S., Damsgaard, J.: A Typology of Multi-Sided Platforms: The Core and the Periphery. In: European Conference on Information Systems, 23, pp. 1–16 (2015)
6. Kane, G.C., Alavi, M., Labianca, G.J., Borgatti, S.: What's different about Social Media Networks? A Framework and Research Agenda. MIS Quarterly, forthcoming (2012)
7. Ellison, N.B.: Social Network Sites: Definition, History, and Scholarship. Journal of Computer - Mediated Communication 13, 210–230 (2007)
8. Buchmann, J.: Internet Privacy. Springer (2012)
9. Shapiro, C., Varian, H.R., Becker, W.E.: Information Rules: A Strategic Guide to the Network Economy. Journal of Economic Education 30, 189–190 (1999)
10. Westin, A.F.: Privacy and Freedom. Atheneum, New York (1967)
11. Westin, A.F.: Social and Political Dimensions of Privacy. Journal of social issues 59, 431–453 (2003)
12. Nolte, C.-G.: Personal Data as Payment Method in SNS and Users' concerning Price Sensitivity - A Survey. In: Business Information Systems Workshops, pp. 273–282 (2015)
13. Schermer, B.W.: The limits of privacy in automated profiling and data mining. Computer Law & Security Review, 45–52 (2011)
14. Stutzman, F., Gross, R., Acquisti, A.: Silent listeners: The evolution of privacy and disclosure on facebook. Journal of privacy and confidentiality 4, 2 (2013)
15. Acquisti, A., Gross, R.: Imagined communities: Awareness, information sharing, and privacy on the Facebook. International workshop on privacy enhancing technologies, 36–58 (2006)
16. Norberg, P.A., Horne, D.R., Horne, D.A.: The privacy paradox: Personal information disclosure intentions versus behaviors. Journal of Consumer Affairs 41, 100–126 (2007)

17. Preibusch, S., Hoser, B., Gürses, S., Berendt, B.: Ubiquitous social networks–opportunities and challenges for privacy-aware user modelling. In: Workshop on Data Mining for User Modelling at UM (2007)
18. Zhang, C., Sun, J., Zhu, X., Fang, Y.: Privacy and security for online social networks: challenges and opportunities. IEEE Network 24, 13–18 (2010)
19. Nolte, C.-G., Brenig, C., Müller, G.: Coherences on Privacy in Social Network Services. A Qualitative System Dynamics Analysis. In: Privacy and Identity Management, 11, forthcoming. Springer, Karlstad, Sweden (2016)
20. Bauer, C., Korunovska, J., Spiekermann, S.: On the value of information-what Facebook users are willing to pay. Ecis 2012 proceedings (2012)
21. Rusthon, K.: DuckDuckGo: The privacy search ruffling Google's feathers, goo.gl/TSH8nj
22. Acquisti, A., John, L.K., Loewenstein, G.: What is privacy worth? The Journal of Legal Studies 42, 249–274 (2013)
23. Tucker, C.: Economics of Privacy and User-Generated Content. Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource (2015)
24. Rochet, J., Tirole, J.: Platform competition in two-sided markets. Journal of the European Economic Association 1, 990–1029 (2003)
25. Armstrong, M.: Competition in two-sided markets. The RAND Journal of Economics 37, 668–691 (2006)
26. Evans, D.S., Schmalensee, R.: Markets with Two-Sided Platforms. Issues in Competition Law and Policy (ABA Section of Antitrust Law) 1 (2008)
27. Hagiu, A., Wright, J.: Multi-sided platforms. International Journal of Industrial Organization 43, 162–174 (2015)
28. Bork, R.H., Sidak, J.G.: What Does the Chicago School Teach About Internet Search and the Antitrust Treatment of Google? Journal of Competition Law and Economics 8, 663–700 (2012)
29. Haucap, J., Kehder, C.: Suchmaschinen zwischen Wettbewerb und Monopol: Der Fall Google. In: Wettbewerb und Regulierung in Medien, Politik und Märkten, pp. 115–154. Nomos Verlagsgesellschaft mbH & Co. KG (2013)
30. Haucap, J., Heimeshoff, U.: Google, Facebook, Amazon, eBay: Is the Internet driving competition or market monopolization? International Economics and Economic Policy 11, 49–61 (2014)
31. Müller, G., Flender, C., Peters, M.: Vertrauensinfrastruktur und Privatheit als ökonomische Fragestellung. Internet Privacy, 143–188 (2012)
32. Knoll, J.: Advertising in social media: a review of empirical evidence. International Journal of Advertising 35, 266–300 (2016)
33. Krasnova, H., Spiekermann, S., Koroleva, K., Hildebrand, T.: Online social networks: why we disclose. Journal of Information Technology 25, 109–125 (2010)
34. Lin, K.-Y., Lu, H.-P.: Why people use social networking sites: An empirical study integrating network externalities and motivation theory. Computers in Human Behavior 27, 1152–1161 (2011)
35. Mankiw, N.G.: Principles of Microeconomics. South-Western Cengage Learning, Mason, Ohio (2012)

36. Kwon, H.E., Oh, W., Kim, T.-H.: One-Sided Competition in Two-Sided Social Platform Markets? An Organizational Ecology Perspective (2015)
37. Zimmermann, C., Nolte, C.-G.: Towards Balancing Privacy and Efficiency. A Principal-Agent Model of Data-Centric Business. International Workshop on Security and Trust Management, 89–104 (2015)
38. Torres, A.M.: Social networking and online privacy: Facebook users' perceptions. Irish Journal of Management (2012)
39. Barary Savadkoohi, F.: Personalized online promotions: long-term Impacts on customer behavior (2012)
40. Tucker, C.E.: Social networks, personalized advertising, and privacy controls. Journal of Marketing Research 51, 546–562 (2014)
41. Ahn, D.-Y., Duan, J.A., Mela, C.F.: An equilibrium model of user generated content. Available at SSRN 1957989 (2011)
42. Mital, M., Sarkar, S.: Multihoming behavior of users in social networking web sites: a theoretical model. Information Technology & People 24, 378–392 (2011)
43. Eisenmann, T., Parker, G., van Alstyne, M.W.: Strategies for two-sided markets. Harvard business review 84, 92 (2006)
44. Facebook introduces new privacy controls. The Irish Times (2008)
45. Gross, D.: Facebook privacy now defaults to friends only, goo.gl/hQZ9ur
46. Prigg, M.: One more reason to Google yourself: Search giant to add privacy information letting users see what it knows about them. Dailymail.com (2016)
47. van Eecke, P.: Technology firms and the European General Data Protection Regulation: How should they prepare?, http://goo.gl/hI7yQJ
48. Vestager, M.: Competition in a Big Data World. Munich, Germany (2016)
49. brt/dpa/AP: Datenschutz-Anpassung: WhatsApp gibt Telefonnummern an Facebook weiter, goo.gl/Pe0sW4
50. Sullivan, D.: Facebook Instant Articles: A Slippery Slope For Google To Do The Same, Hurting The Web?, goo.gl/ui0PZz
51. Constine, J.: Wait, Did Facebook Just Build A Kickstarter Competitor?, goo.gl/Eagrp7
52. Kaiser, U., Wright, J.: Price structure in two-sided markets: Evidence from the magazine industry. International Journal of Industrial Organization 24, 1–28 (2006)
53. Athey, S., Calvano, E., Gans, J.S.: The Impact of Consumer Multi-Homing on Advertising Markets and Media Competition (2012)
54. Olbrich, R., Holsing, C.: Facebook Ads. WIST 43, 557–560 (2014)
55. Duggan, M., Ellison, N.B., Lampe, C., Lenhart, A., Madden, M.: Social Media Update 2014 (2015)
56. eMarketer: Net digital advertising revenue share of major ad-selling online companies worldwide from 2012 to 2014 (2014)
57. DeLong, J.B., Summers, L.H.: The'new economy': background, historical perspective, questions, and speculations. Economic Review-Federal Reserve Bank of Kansas City 86, 29 (2001)

Threat Intelligence Sharing Platforms: An Exploratory Study of Software Vendors and Research Perspectives

Clemens Sauerwein[1], Christian Sillaber[1], Andrea Mussmann[1,], and Ruth Breu[1]

[1] University of Innsbruck, Department of Computer Science, Innsbruck, Austria
{Clemens.Sauerwein,Christian.Sillaber,Andrea.Mussmann,Ruth.Breu}
@uibk.ac.at

Abstract. In the last couple of years, organizations have demonstrated an increased willingness to exchange information and knowledge regarding vulnerabilities, threats, incidents and mitigation strategies in order to collectively protect against today's sophisticated cyberattacks. As a reaction to this trend, software vendors started to create offerings that facilitate this exchange and appear under the umbrella term "Threat Intelligence Sharing Platforms". To which extent these platforms provide the needed means for exchange and information sharing remains unclear as they lack a common definition, innovation in this area is mostly driven by vendors and empirical research is rare. To close this gap, we examine the state-of-the-art software vendor landscape of these platforms, identify gaps and present arising research perspectives. Therefore, we conducted a systematic study of 22 threat intelligence sharing platforms and compared them. We derived eight key findings and discuss how existing gaps should be addressed by future research.

Keywords: Information Security, Threat Intelligence Sharing Platform, Information and Knowledge Sharing

1 Introduction

As information systems used in organizations today are characterized by continuously increased complexity, heterogeneity and interconnectedness, the number and complexity of security related incidents increases accordingly [1, 2]. Recent prominent security incidents have shown the growing spectrum of possible attacks and the immense business harm that result from these attacks, including a devastating loss of intellectual property, productivity, money and reputation [1, 3].

Therefore, the protection of an organization's infrastructure, vulnerability management, internal dissemination of cyber security information and security training have become key activities [4]. Beside these traditional countermeasures, there is an observable trend to increasingly exchange information and knowledge between trusted organizations to aid the management of vulnerabilities and threats and to mitigate incidents [5, 6]. To support these efforts, several standards (e.g. CybOX, STIX, or TAXII) have been developed to enable the automated exchange of threat intelligence [7, 8, 9, 10].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Sauerwein, C.; Sillaber, C.; Mussmann, A.; Breu, R. (2017): Threat Intelligence Sharing Platforms: An Exploratory Study of Software Vendors and Research Perspectives, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 837-851

In 2013 Dandurand and Serrano introduced the concept of a threat intelligence sharing platform, whose requirements are to: (a) Facilitate information sharing,

(b) enable automation, and (c) facilitate the generation, refinement and vetting of data [11]. Today's threat intelligence sharing solutions more or less build on these requirements and several software vendors are offering products [12]. While there are already a variety of solutions on the market, the majority of publications investigate fundamental requirements and challenges for the development of threat intelligence sharing platforms [11, 13, 12, 14].

To which extent these platforms provide the needed means for threat intelligence sharing remains unclear since no scientific analysis of the state-of-the-art of threat intelligence sharing platforms exists and no empirical research has been conducted yet. To address this gap, we seek to answer the following research questions: (a) *What is the state-of-the-art of threat intelligence sharing platforms?* (b) *What are existing gaps in currently available threat intelligence sharing platforms?* and (c) *What are the implications for scientific research in this area?*

The goal of our research is therefore to provide a comprehensive analysis and comparison of the threat intelligence sharing platform market. Based on our findings we discuss the implications for scientific research and their significance.

To achieve this goal, we conducted a systematic study of 22 threat intelligence platforms, consisting of open and closed source products. We compared these solutions according to different dimensions, such as covered use cases, intelligence sharing functionalities and collaboration capabilities. Based on these results, we derived eight key findings and discuss their implications for scientific research.

Our research has shown that, although the interest in this domain has considerably increased over the past years, a common definition of threat intelligence sharing platforms is still missing. While STIX is the extensive de-facto standard for describing threat intelligence data, most platforms do not fully utilize its descriptive capabilities. This is illustrated by the fact that the majority of platforms primarily focus on the sharing of indicators of compromise.

The remainder of this paper is structured as follows. Section 2 provides related work regarding threat intelligence sharing platforms and related studies. Section 3 outlines the underlying research methodology and procedure carried out. Section 4 discusses the key findings. Section 5 discusses the results and implications for scientific research. Finally, Section 6 concludes the paper and provides outlook on future research.

2 Related Work

In the past, organizations used ad-hoc solutions such as email messages, phone calls, ticket systems, or face-to-face meetings to communicate security related information. Recently, a trend to form interconnected communities of exchange via associated platforms for the (semi-) automated exchange of security related threat intelligence can be observed [12]. Moreover, the exchange of security knowledge between experts across organizations presents a potential countermeasure to protect against today's sophisticated threat actors, as not every organization has the resources to develop an adequate information security program independently, and organizations can benefit from other organizations' experiences and knowledge [5, 15].

Several researchers identified challenges and requirements for threat intelligence sharing platforms [11, 13, 12, 14]. In addition, several standardization efforts have been made to facilitate cyber security data sharing in a standardized manner, e.g. Open Incident of Compromise(OpenIOC), Cyber Observable eXpression (CybOX), Structured Threat Information eXpression (STIX), Incident Object Description Exchange Format (IODEF) or Trusted Automated eXchange of Indicator Information (TAXII) [8, 9, 10]. Moreover these standards form more or less the basis for today's threat intelligence sharing platforms. In order to select the right standard for a particular use case Burger et al. provide an agnostic framework in which standards can be evaluated and assessed [16].

Only a few studies and descriptions of existing threat intelligence sharing platforms can be identified. In [17] the SANS Institute gives an overview of a small selection of open source threat intelligence platforms, including the Collective Intelligence Framework (CIF), Collaborative Research into Threats (CRITs), MANTIS Cyber-Intelligence Management Framework, Malware Information Sharing Platform (MISP), and Soltra Edge, and conclude that the market for threat intelligence sharing is still developing. Brown et al. discuss the community requirements and expectations of an all-encompassing threat intelligence management platform based on studies on a few threat intelligence sharing platforms [12]. These contributions take into consideration only a subset of available threat intelligence sharing platforms without providing a comprehensive analysis of the state-of-the-art. Several tool and vendor studies loosely related to threat intelligence sharing in organizations can be found, e.g. on software vendors in the area of Governance, Risk & Compliance (GRC) management software [18], cloud providers [19] and the security software market [20].

Moreover, as part of our research methodology we use a multivocal literature review [21]. Only few systematic literature reviews using this methodology can be found in information systems research [22, 23, 24]. As outlined by Garousi et al., multivocal literature reviews can be valuable in closing the gap between academic research and practice [25].

To the best of our knowledge, no prior research has been conducted on threat intelligence sharing platforms in order to provide a comprehensive analysis of existing software solutions and research perspectives that could guide research and industry.

3 Research Methodology

Our research is based on an exploratory study carried out in close collaboration with a security expert group developing a nation-wide threat intelligence sharing community in a German-speaking country. It consisted of two parts and was conducted between May 2016 and July 2016. At first two workshops with the community were held to form the basis of our subsequent multivocal literature review on threat intelligence sharing platforms in industrial and academic context. In order to guarantee reproducibility of the applied research methodology, a research protocol was developed, including workshops transcripts, search term definitions, search results, platform selections, data extractions and platform classifications.

3.1 Workshops

Two focus group discussions [26] with respectively ten security experts from industry being members of the aforementioned security expert group were conducted. Table 1 gives an overview of the participants, their organizational roles, qualifications, type and size of their organizations. The findings from the first round were evaluated and refined in the second round. The primary goal of these two workshops was to get an understanding what the participants recognize as threat intelligence sharing platform and to compile a list of platforms they either use or consider for usage at their organization.

Table 1. Overview Workshop Participants

ID	Organizational Role	Qualifications		Type of Org.	# of Employees
		Security Specific Certifications	*University Degree in CS or IS*		
1	Did not disclose	CISSP	x	Finance	>1000
2	Security Operations Team Leader		x		<150
3	Security Analyst			Fiance	>1000
4	Security Operations Officer	CISSP, CCSP		Insurance	>1000
5	IT Security Architect		x	IT	150 - 1000
6	IT Security Analyst	CISSP, CEH	x	Finance	>1000
7	Cyber Security Incident Response Team			Finance	>1000
8	Managed Security Service Provider		x		>1000
9	Did not disclose		x	Education	<150
10	Security Specialist		x	Finance	150 - 1000
11	Information Security Officer	CISM		Finance	>1000
12	Cyber Security Task Force	CISSP, CISA	x	Finance	150 - 1000
13	Head of Security Operations Center	CISA	x	Finance	>1000
14	Security Consultant	CISM, CISA	x	Consulting	>1000
15	Security	CISM, CISSP	x	Finance	>1000
16	Cyber Security Officer	CISM, CISA	x	Finance	>1000
17	Cyber Security Incident Response Team	CISSP, CEH	x	Production	150 - 1000
18	Did not disclose	CISSP	x	Finance	>1000
19	Did not disclose				
20	Did not disclose				

The workshops were held in parallel after the participants were instructed by two researchers. The aim of these parallel workshops was to limit bias from other participants and to get more comprehensive results. The researchers asked questions during the discussions to clarify participant’s statements. Each workshop was recorded and then transcribed. Finally, qualitative summaries were produced [27] from the discussions in order to derive keywords and a common understanding for the subsequent multivocal literature review. In addition, a comprehensive list of threat intelligence sharing platforms used by participants was compiled to evaluate the results of the systematic search.

3.2 Multivocal Literature Review

Secondly, we conducted a multivocal literature review (MLR) to identify relevant threat intelligence sharing platforms used in research and practice through a systematic analysis of academic literature and grey literature (e.g. blogs, white papers, webpages) [21, 25]. Moreover, a MLR closes the gap between research and practice [28] and provide a more comprehensive picture on the state-of-the-art of a particular research field [25]. According to that there are no systematic guidelines for conducting MLRs in computer science [25], we oriented us on [23] which conducted a MLR in information system research. In doing so, we derived the following three research steps:

Search Strategy: At first we conducted a systematic search where we used the following academic search engines as well as ordinary web search engines: ACM Digital Library, Cite Seer, ScienceDirect, Google Scholar, IEEE Digital Library, Springer, Scopus, Taylor&Francis and Wiley, Google and Bing. According due that there isn't a common definition and consistent description for the approach of exchanging threat intelligence we derived the following keywords from the discussions during the two workshops: *(Cyber Security OR Threat) AND (Intelligence OR Information OR Data) AND Sharing (Platform OR Tool)* We used these keywords for our search where we tried to identify threat intelligence sharing platforms through analysis of the titles and reading the abstracts of the obtained academic and grey literature. In doing so, we obtained a list of 31 threat intelligence sharing platforms with corresponding references including academic literature as well as grey literature (e.g. white papers, vendor specific web pages). We compared the resulting list of threat intelligence sharing platforms with the list of platforms compiled by the participants of the workshops and with commercial market studies carried out by Gartner [29] and Forrester [30]. The comparison resulted in an unchanged list of 31 threat intelligence sharing platforms for further investigations.

Inclusion and Exclusion: Based on the list of 31 threat intelligence sharing platforms we collected for each platform as many artifacts as possible. These artifacts included websites, white papers, discussions on blogs/forums, technical reports, scientific papers and tool demos. According to that there are apparent issues of reliability and validity associated with grey literature [23] we ranked the quality of the artifacts with respect to the trustworthiness and reliability of the sources. Based on the compiled list of platforms with corresponding references, we applied the following selection procedure: Based on reading of the identified artifacts, we excluded tools that either do not enable the sharing of threat intelligence or do not address organizational cyber security (such as general purpose wikis). We included tools, when the evaluated quality of artifacts was adequate (e.g. websites provided more than a couple of buzzwords), they were published since 2010, and dealing with (cyber security-) threat intelligence sharing. This procedure resulted in a final set of 22 threat intelligence sharing platforms for further analysis.

Classification of platforms: Finally, we analyzed these 22 platforms. At first, we assessed the platforms' web pages and studied the provided documentations. Secondly, if there was a free or trial-version available, the software was tested locally. Since this was not possible for every threat intelligence sharing platform, we also analyzed all product videos that were available. During our gradual approach we analyzed every platform according to the following perspectives: Licensing model, use cases, supported standards, supported threat intelligence constructs, shared information/threat intelligence, sharing functionalities, collaboration capabilities, integration capabilities, analysis, deployment, and provided user interfaces.

4 Study Results and Key Findings

The applied methodology, described in the previous Section identified the following 22 threat intelligence sharing platforms: Accenture Cyber Intelligence Platform, Anomali ThreatStream, Anubis Networks Cyberfeed, BrightPoint Security Sentinel, Collaborative Research into Threats (CRITs), Comilion, Facebook Threat Exchange, Falcon Intelligence Crowdstrike, MANTIS Cyber Threat Intelligence Management Framework, Malware Information Sharing Platform (MISP), McAfee Threat Intelligence Exchange, Microsoft Interflow, Open Threat Exchange (OTX), Soltra Edge, HP ThreatCentral, ThreatCloud IntelliStore, ThreatConnect, ThreatQ, ThreatTrack ThreatIQ, Eclectic IQ, IBM X-Force Exchange, Collective Intelligence Framework (CIF).

The classification and analysis of the identified platforms resulted in eight key findings. They are described in this chapter and examples are given. A summary of the tool classification is also provided in Table 2.

Table 2. Overview analysed tools

Type of platform	**#**	**Supported standards**	**#**
Sharing of threat intelligence	8	STIX	10
Sharing of aggregated data	7	OpenIOC	2
Security information repository	4	STIX & OpenIOC	2
Other	3	IODEF	1
		Other (proprietary)	7
Shared information focuses on		**Licensing model**	
Indicators of Compromise	10	Closed Source (comm.)	16
All types of STIX's constructs	4	Open Source	4
Other	8	Free-to-use	2

4.1 Key Finding 1: There is no common definition of threat intelligence sharing platforms

Beside the standards for describing (e.g. STIX) and sharing (e.g. TAXII) of threat intelligence, research and practice have not yet developed a comprehensive definition and common understanding of what constitutes a threat intelligence sharing platform. Therefore, different types of platforms were identified:

Eight of the identified platforms focus on the sharing of threat intelligence between organizations. While they aggregate information from the users participating in the platform, seven platforms only share only data (and not intelligence in its strictest sense) that is automatically aggregated from various available paid and open information security data sources (cf. Open Source Intelligence). One of the identified platforms provides a hybrid form of a threat intelligence sharing platform, where data from multiple data sources is combined with the threat intelligence provided by participating users. Moreover, there are four tools which only consist of a central repository which provides context specific security information (e.g. information about malware). The two remaining platforms focus on the sharing of technical data (e.g. SNORT rules) between security applications.

Due to this heterogeneity and diversity of threat intelligence platforms, prospective end users are challenged with the selection of a platform for their particular use case.

4.2 Key Finding 2: STIX is the de-facto standard for describing threat intelligence

The landscape of standards available to describe threat intelligence is rather small compared to the number of sharing platforms available. Our analysis showed that most threat intelligence sharing platforms rely on standards such as OpenIOC, STIX, and IODEF. More than two-thirds of the analyzed platforms provide direct import and export capabilities supporting the standards mentioned above. In detail, ten platforms rely on STIX, two on OpenIOC, two on both of them, and one platform on IODEF. For example, the Open Threat Exchange (OTX) platform provides STIX as well as OpenIOC import and export functionalities.

We found that STIX is the most commonly used standard and can be considered as the de-facto standard for describing threat intelligence. It builds upon the CybOX, CAPEC, MAEC and CVRF standards, and provides a unifying architecture tying together a diverse set of cyber threat information [31]. The STIX architecture consists of eight core cyber threat concepts as independent and reusable constructs and takes their interrelationship into account. The eight constructs describing Cyber Observables (e.g. IP addresses, file names, hashes), Indicators, Incidents, Adversary Tactics Techniques and Procedures (including attack patterns, kill chains, etc.), Exploit Targets (e.g. vulnerabilities, weaknesses), Courses of Action (e.g., incident response, mitigation strategies), Cyber Attack Campaigns, and Cyber Threat Actors. These constructs can be - at least partially - found in all analyzed platforms.

Moreover, these constructs can be used to provide meaningful inputs to information security processes like prevention, detection, or response. For example, valuable inputs for response processes are shared Course of Actions with corresponding Incident descriptions.

4.3 Key Finding 3: Platforms primarily focus on sharing of indicators of compromise

The observed platforms primarily focus on the sharing of indicators of com- promise, e.g. the Open Threat Exchange (OTX) platform. Indicator of compromise include information that enable the identification of potentially malicious activities. For example, indicators of compromise are malicious IP addresses, anomalous user activities, descriptions of malicious files, etc. While the OpenIOC standard is primarily designed to share them, the analyzed platforms use the STIX's Observable and

Indicator constructs to describe them. Although the STIX standard is rather expressive (cf. Key Finding 2), the majority of platforms use the two aforementioned STIX constructs.

4.4 Key Finding 4: The Majority of platforms is closed source

There are six free- to-use threat intelligence sharing platforms available on the market, whereof four are open source tools released under the GNU General Public License, including the Malware Information Sharing Platform (MISP), Collective Intelligence Frame- work (CIF), Collaborative Research Into Threats (CRITs) and MANTIS Cyber- Intelligence Management Framework. The Open Threat Exchange (OTX) and Soltra Edge platform are free-to-use but were not released under an open source license. The remaining 16 are closed source.

4.5 Key Finding 5: Most platforms focus on data collection instead of analysis

The "intelligence" provided by the majority of threat intelligence sharing platforms does not constitute "intelligence" in the traditional sense. In the context of information security "intelligence" is the product of the intelligence lifecycle model, which includes several activities like planning, data collection, analysis and, dissemination [32, 33].

However, we found that the majority of tools primarily focuses on data collection and more or less neglects the other activities of the intelligence lifecycle. Therefore, most currently available threat intelligence platforms resemble data warehouses more than "real" intelligence sharing platforms.

Moreover, they provide limited analysis and visualization capabilities and lag behind comparable knowledge sharing platforms and data mining solutions from other domains. This is insofar surprising as the value of these platforms is constrained by the user's ability to interpret, absorb, enhance and react to the provided information [34]. Moreover, only a few platforms provide interfaces for third party tools that would enable further analysis of the received threat information.

Threat intelligence sharing platforms currently provide basic analysis capabilities, such as browsing, attribute based filtering and searching of information. Additionally, only a small fraction of platforms implement pivot functionalities which enable the visualization of relationships between the threat intelligence constructs.

4.6 Key Finding 6: Trust issues between users and platform providers are mostly neglected

Our research showed that there are two possible perspectives on trust: the organization, which uses such a platform, towards the provider and vice-versa. Since organizations may share private or sensitive information it is necessary to establish a trust-bond between organizations and the provider of said threat intelligence sharing platform. Moreover, the provider or other organizations must be able to trust the information provided by a particular organization. This means that users of a threat intelligence sharing platform must be careful when dealing with intelligence provided by the platform. In addition, the access to these platforms must be restricted in order to avoid any unauthorized access. For example, the shared intelligence might be of potential interest to attackers.

In order to overcome these trust and access control issues, threat intelligence sharing platforms must provide control mechanisms to specify what information is shared, how much of it and with whom. In addition, access control plays an important role to these platform, since these platforms might be of potential interest to attackers.

Accordingly, six platforms provide trust modelling functionalities, like forming trusted and closed communities, peer-to-peer connections, anonymization of shared threat intelligence, or policies for maximum control of privacy and security. However, they are mostly limited to group-based access control and ranking mechanisms.

Moreover, this topic gained traction in research as well. For example, Steinberger et al. present a trust model that determines trust ratings for security events [35]. Murdoch and Leaver discuss the barriers to participate in a threat intelligence sharing platform caused by the conflict between the need for anonymity versus the need to trust the shared information [36]. In [37] the authors introduce privacy principles for sharing cyber security data that can reduce the risk of data exposure and help to manage trust.

4.7 Key Finding 7: Academic and commercial interest in threat intelligence sharing increases

In November 2011, the OpenIOC standard was released and laid the foundation for threat intelligence sharing. Between 2010 and 2012 only little attention was drawn to threat intelligence sharing in research and practice. In 2013 Dandurand and Serrano introduced the first concept of a threat intelligence management platform with associated requirements [11]. Between 2013 and 2014 the comprehensive STIX and TAXII standards were released. Since then, the number of publications and vendors providing threat intelligence sharing platforms has grown remarkably. For example, the total amount of publications in 2015 was more than threefold compared to 2014. As the market for threat intelligence sharing platforms is relatively new and still developing [17], it can be expected that the number of platforms and scientific publications will continue to grow in the near future.

4.8 Key Finding 8: Many manual tasks make the user the bottleneck

Threat intelligence sharing platforms provide limited automated data integration capabilities. Therefore, a lot of manual user interaction for sharing and acquiring valuable intelligence is necessary. Moreover, the success of a threat intelligence platform depends on the willingness of users to share intelligence which is limited by the organizations' availability of free resources and the employee's motivation to actively participate. As most platforms lack automated means of intelligence gathering, and more importantly, automatic sanitation of sensitive intelligence, these activities still require manual effort. Beside classical file importing functionalities, most threat intelligence sharing platforms lack convenient user interfaces for quickly adding new data records and require many user interactions to achieve the desired goal.

5 Discussion and Research Implications

In this section we provide a discussion on the limitations of this exploratory study and a discussion on the results and their implications for research.

5.1 Limitations

Our exploratory study might be limited by certain threats to validity. Limitations that have to be acknowledged and accounted for are an (i) incomplete list of threat intelligence sharing platforms, (ii) wrong definition of key words and inclusion/exclusion criteria, (iii) wrong classification and analysis of platforms, and (iv) incomplete or biased descriptions of platforms.

In order to counteract (i), we evaluated the list of identified platforms with the results of our stakeholder workshops and commercial market studies on threat intelligence sharing platforms. However, there might be the possibility of non-identifiable tools if they were released after our cutoff day (after July 2016) or there is not any public information available about them.

There might be the possibility of type (ii) limitations. In order to avoid them the keywords and inclusion/exclusion criteria were developed based on expert opinions collected during our workshops with the stakeholders.

To inhibit (iii), we have chosen a type of cross-validation approach in which each contributor to this research was given a subset of platforms to analyze and classify that intersected with another contributor's set. Hence, classification discrepancies were discovered and limited through re-classification and analysis. Finally, to overcome (iv), we tried to use more than one information source to classify and analyze a platform.

5.2 Discussion of results and implications for research

Our comprehensive exploratory study pointed out that there is an increasing interest in threat intelligence sharing in research and practice, i.e. the number of publications that apply to his research and number of platforms increased over the last three years. Moreover, it seems that there is a different focus in research and practice since several publications discuss the principles of threat intelligence sharing although there are already a variety of solutions on the market. This might be traced back to a missing common understanding of what threat intelligence sharing is about, due to the diversity of threat intelligence sharing platforms. Hence, one of the biggest gaps is the lack of a common definition and characterization of threat intelligence sharing platforms.

Aware of the possible limitations of the research at hand, the following implications for research, based on the eight key findings, can be derived:

Since key finding 1 and 5 showed that software vendors have a different understanding on threat intelligence sharing, it is necessary to develop a standardized definition and characterization of threat intelligence sharing platforms. In this context it might be beneficial to adopt the wide spread intelligence life cycle model, including planning, collection, analysis, and dissemination activities, to the threat intelligence sharing domain in order to generate intelligence. Thereby, it might be necessary to investigate and define how the different activities within the model can be addressed by a threat intelligence sharing platform. Moreover, these standardization efforts might pave the way for a prospective threat intelligence sharing platform which provides "real" intelligence instead of data warehousing and limited data analysis capabilities. Additionally, organizations might benefit from a common understanding as well, since it might simplify the selection of an appropriate threat intelligence platform.

Key finding 2 showed that three standards are used to facilitate the description of threat intelligence of which STIX is the most used. We believe that it is becoming the de-facto standard in the field. STIX is a detailed and extensive standard consisting of eight constructs which enable the description of a broad range of security related information and their relationships. While the number of standards and available exchange formats is limited at the moment, a trend towards use case specific description formats can be observed (e.g. internal sharing vs sharing across the organizational boundaries).

According to key finding 3 the majority of tools only share indicators of compromise which can be described by two constructs of the STIX standard. Based on this observation the following two implications can be derived: (a) Standards for describing threat intelligence are too generic and powerful, or (b) only the low hanging fruits, namely indicators of compromise, are shared at the moment. In order to get deeper insights on this issue, empirical research on the expected, needed and shared information within a threat intelligence sharing platform is needed.

One argument for using a threat intelligence sharing platform is resource reduction through sharing security knowledge and information. However, as outlined in key finding 5, the majority of tools are rather data-warehouses than intelligence sharing platforms. Consequently, organizations must often evaluate the received information which might result in a lot of additional work. In order to address this issue, research in this area should focus on moving away from mere security data sharing towards knowledge and ultimately intelligence sharing.

As key finding 5 showed that threat intelligence sharing platforms suffer from deficient analysis and visualization functionalities and key finding 8 showed that the submission of new threat intelligence is hindered by limited input options, existing platforms and their user interfaces should be scientifically evaluated to identify potential weaknesses and requirements. In doing so, empirical studies on the required functionalities and visualization options of threat intelligence sharing platforms should be conducted.

Key finding 6 showed that trust plays a paramount role in the context of threat intelligence sharing platforms. To some extent threat intelligence platforms already provide preliminary functionalities to establish trust between the collaborators. In order to support the ongoing research and provide a generally accepted model to guarantee trust, empirical research on the sharing behavior of users and their expectations on data privacy and security is needed.

Key finding 7 showed that there is an increasing interest of threat intelligence sharing in research and practice. Additionally, key finding 4 states that the majority of platforms are closed source. Accordingly, there might be a lack of open threat intelligence platforms and open data sets for scientific research, e.g. to conduct empirical studies.

In order to address this gap it might be necessary that research collaborates with industry. Furthermore, such collaborations imply new research questions resulting from arising problems in practice, and provide potential for future research.

In order to address key finding 8 it is necessary to conduct empirical research on how to motivate users and organization to share information on a threat intelligence sharing platform. For example, it might be necessary to develop and evaluate incentive mechanisms to foster the collaboration within threat intelligence sharing platforms.

Our objective is to explore potential research opportunities by pointing to research questions that (1) investigate threat intelligence as an artifact; (2) see threat intelligence data within an ecosystem of competing and complementary frameworks and standards; or (3) evaluate the usage of threat intelligence in organizations.

Threat intelligence as an artifact: There is a clear need to investigate the foundations, design, applicability, and internal consistency (or lack thereof) in threat intelligence sharing formats and standards. For example, the dominating STIX standard includes many significant artifacts that can be used to describe different aspects of cyber threats (e.g. Indicators of Compromise, Techniques-Tactics & Procedures). However, the success of this expressiveness is not yet clear. Furthermore, little is known about the quality requirements for threat intelligence data artifacts [38]. Questions include: (1) *How can the expressiveness of threat intelligence exchange formats be compared?* (2) *Which concepts of threat intelligence formats are superficial?*

Threat intelligence ecosystem: The core principle of the design of most standards for threat intelligence exchange is to align systematically with cognate standards and formats. These include specialized formats (e.g. IOC, CybOX) as well as formats of higher abstraction (e.g. STIX, TAXII). Understanding how threat intelligence analysis operates in an ecosystem of competing and collaborating standards and frameworks is important. Questions include: *(3) How can organizations manage the integration of different threat intelligence data sources?* (4) *How can organizations decide on which data sources to include?* (5) *How can organizations identify missing data sources?* (6) *Are fewer but complex formats better than many, diversified formats?*

Threat intelligence in use: Little is known on the actual value provided to organizations that participate in threat intelligence sharing platform. There is no academic research about the value proposition of threat intelligence sharing platforms. Questions include: (7) *How can an organization evaluate the impact of threat intelligence sharing platforms?* (8) *How is threat intelligence disseminated within organizations?* (9) *What is the impact of threat intelligence sharing on organizational decision making processes?* (10) *How can organizations be motivated to participate in threat intelligence sharing platforms?* (11) *How can trust be established in threat intelligence sharing platforms?* (12) *How can stakeholders be incentivized to participate in threat intelligence sharing platforms?* (13) *Which visualization and query options are required by organizational stakeholders?* (14) *Which impact do different levels of participation have in threat intelligence sharing platforms?*

6 Conclusion

In this paper we presented an exploratory study on software vendors of threat intelligence sharing platforms and derived future research perspectives. Therefore, we conducted two workshops and a Multivocal Literature Review, including academic and grey literature. It identified a list of 22 threat intelligence sharing platforms used in research and practice. With respect to our research questions we briefly analyzed them and elicited eight key findings. For example, our key findings identified that there is an increasing interest towards threat intelligence sharing, research and practice lacks a consistent definition of threat intelligence sharing, and current threat intelligence sharing is comparable to data warehousing and doesn't provide "real" intelligence. Based on the key findings, we discussed several implications for future research focusing on the creation of a common understanding of threat intelligence sharing and the improvement of current practice. Our future work will focus on empirical research in order to provide a common definition and characterization of threat intelligence sharing platforms for research and practice.

Acknowledgment

This work was partially supported by the Austrian Federal Ministry of Science, Research and Economics (BMWFW), FFG Project 855383 SALSA (ICT of the Future).

References

1. Marinos, L., Belmonte, A., Rekleitis, E.: ENISA Threat Landscape 2015. Technical Report, ENISA - The European Union Agency for Network and Information Security (2016)
2. Miller, A., Horne, R., and Porter, C.: 2015 information security breaches survey. Technical Report, PWC- PrisewaterhouseCoopers (2015)
3. PWC: The Global State of Information Security Survey 2016. Technical Report, PWC-PrisewaterhouseCoopers (2016)
4. Puhakainen, P., Siponen, M.: Improving employees' compliance through information systems security training: an action research study. Mis Quarterly, 757–778 (2010)
5. Fenz, S., Heurix, J., Neubauer, T., Pechstein, F.: Current challenges in information security risk management. J. Information Management & Computer Security 22, 410–430 (2014)
6. Fransen, F., Smulders, A., Kerkdijk, R.: Cyber security information exchange to gain insight into the effects of cyber threats and incidents. J. e & i Elektrotechnik und Informationstechnik 132, 106–112 (2015)
7. Kert, M., Lopez, J., Evangelos, M., Preneel, B.: State-of-the-Art of Secure ICT Landscape. Technical Report, ENISA - NIS Platform - Working Group 3 (2014)
8. Kampanakis, P.: Security automation and threat information-sharing options. J. IEEE Security & Privacy. 12, 42–51 (2014)
9. Steinberger, J., Sperotto, A., Golling, M., and Baier, H.: How to exchange security events? overview and evaluation of formats and protocols. In: IEEE International Symposium on Integrated Network Management (IM), pp. 261–269. IEEE, Los Almitos (2015)

10. Martin, R.A.: Making security measurable and manageable. In: IEEE Military Communications Conference, pp. 1–9. IEEE, Los Almitos (2008)
11. Dandurand, L., Serrano, O.: Towards improved cyber security information sharing. In: 5th International Conference on Cyber Conflict (CyCon), pp. 1–16. IEEE, Los Almitos (2013)
12. Brown, S., Gommers, J., Serrano, O.: From Cyber Security Information Sharing to Threat Management. In: 2nd ACM Workshop on Information Sharing and Collaborative Security, pp. 43–49. ACM, New York (2015)
13. Serrano, O., Dandurand, L., Brown, S.: On the design of a cyber security data sharing system. In: ACM Workshop on Information Sharing and Collaborative Security, pp. 61–69. ACM, New York (2014)
14. Appala, S., Cam-Winget, N., McGrew, D., Verma, J.: An Actionable Threat Intelligence system using a Publish-Subscribe communications model. In: 2nd ACM Workshop on Information Sharing and Collaborative Security. ACM, New York (2015)
15. Ernest Chang, S., Lin, C.-S.: Exploring organizational culture for information security management. J. Industrial Management & Data Systems 107, 438–458 (2007)
16. Burger, E. W., Goodman, M. D., Kampanakis P., Zhu, K.A.: Taxonomy model for cyber threat intelligence information exchange technologies. In: ACM Workshop on Information Sharing & Collaborative Security, pp. 51–60. ACM, New York (2014)
17. Poputa-Clean, P.: Automated Defense - Using Threat Intelligence to Augment Security. Technical Report, SANS Institute InfoSec (2015)
18. Racz, N., Weippl, E., Seufert, A.: Governance, risk & compliance (GRC) software-an exploratory study of software vendor and market research perspectives. In: 44th Hawaii International Conference on Information Sciences (HICSS), pp. 1–10. IEEE, Los Almitos (2011)
19. Repschlaeger, J.: Transparency in cloud business: Cluster analysis of software as a service characteristics. In: International Conference on Grid and Pervasive Computing, pp. 1–10. Springer, Berlin Heidelberg (2013)
20. Dey, D., Lahiri, A., Zhang, G.: Quality Competition and Market Segmentation in the Security Software Market. Mis Quarterly 38, 589–606 (2014)
21. Ogawa, R. T., Malen, B.: Towards rigor in reviews of multivocal literatures: Applying the exploratory case study method. J. Review of Educational Research. 61/3, 265–286, (1991)
22. Ampatzoglou, A., Chatzigeorgiou, A., Avgeriou, P.: The financial aspect of managing technical debt: A systematic literature review. J. Information and Software Technology. 64, 52–73 (2015)
23. Tom, E., Aurum, A., Vidgen, R.: An exploration of technical debt. Journal of Systems and Software. 86/6, 1498–1516 (2013)
24. Sulayman, M., Mendes, E.: A systematic literature review of software process improvement in small and medium web companies. In: International Conference on Advanced Software Engineering and Its Applications, pp. 1-8. Springer (2009)
25. Garousi, V., Felderer, M., Mantyla, M.V.: The need for multivocal literature reviews in software engineering: complementing systematic literature reviews with grey literature. In.: 20th International Conference on Evaluation and Assessment in Software Engineering, p. 26. ACM, New York (2016)
26. Wilkinson, S.: 10 focus group reseach. J. Qualitative research: Theory, method and practice, p. 177 (2004)
27. Mayring P., Gläser-Zikuda, M.: Die Praxis der Qualitativen Inhaltsanalyse. Beltz, Weinheim, (2008)

28. Elmore, R.F.: Comment on towards rigor in reviews of multivocal literatures: Applying the exploratory case study method. J. Review of Educational Research. 61, no. 3, pp. 293–297 (1991)
29. McMilian, R., Pratap, K.: Market Guide for Security Threat Intelligence Services. Technical Report, Gartner (2014)
30. Holland, R., Balauras, S., Blackborow, J.: The State of the Cyberthreat Intelligence Market. Technical Report, Forrester (2015)
31. Barnum, S.: Standardizing cyber threat intelligence information with the Structured Threat Information eXpression (STIX™). Technical Report, MITRE Cooperation (2012)
32. Gill, P., Phythian, M.: Intelligence in an insecure world. Polity, Cambridge (2006)
33. Heuer, R.J.: Psychology of intelligence analysis. Technical Report, CIA (1999)
34. Sander, T., Hailpern, J.: Ux aspects of threat information sharing platforms: An examination & lessons learned using personas. In: 2nd ACM Workshop on Information Sharing and Collaborative Security, pp. 51-59. ACM, New York (2015)
35. Steinberger J., Kuhnert, B., Sperotto, A., Baier, H., Pras, A.: In whom do we trust-sharing security events. In: International Conference on Autonomous Infrastructure, Management and Security, pp. 111-124. Springer, Berlin (2016)
36. Murdoch, S., Leaver, N.: Anonymity vs. trust in cybersecurity collaboration. In: 2nd ACM Workshop on Information Sharing and Collaborative Security, pp. 27-29, ACM, New York (2015)
37. Fisk, G., Ardi, C., Pickett, N., Heidemann, J., Fisk, M., Papadopoulos, C.: Privacy principles for sharing cyber security data. In: IEEE Security and Privacy Workshops (SPW), pp.193-197, IEEE, Los Almitos (2015)
38. Sillaber, C., Sauerwein, C., Mussmann, A., Breu, R.: Data quality challenges and future research directions in threat intelligence sharing practice. In: 3rd ACM Workshop on Information Sharing and Collaborative Security, pp.65-70 ACM, New York (2016)

Conceptualization of Relational Assurance Mechanisms – A Literature Review on Relational Assurance Mechanisms, Their Antecedents and Effects

Michael Lang[1], Manuel Wiesche[1], Helmut Krcmar[1]

[1] Technical University of Munich, Department of Informatics, Munich, Germany
{michael.lang,wiesche,krcmar}@in.tum.de

Abstract. Assurance mechanisms are an important element of relational governance and frequently used in information systems (IS) research; still missing in this field, however, is a coherent and interrelated structure to organize available knowledge. In this study, we provide a first step towards development of a conceptualization framework of relational assurance mechanisms to enable their further investigation. From our analysis of existing literature, we discover two gaps in assurance research: (1) a fragmentation of assurance research and (2) a lack of conceptual consensus on relational assurance mechanisms. We provide a theoretical framework consisting of a conceptualization of identified relational assurance mechanisms, their antecedents and effects as a means of advancing theory in this area. Several possibilities for future research are discussed.

Keywords: relational governance, relational assurance mechanism, conceptualization, psychological control perspective, literature review

1 Introduction

In recent years, relational governance of inter-organizational relationships has emerged as a dominant perspective in exchange relationships [1]. Within information systems (IS) research, attention has been focused on how relational governance complements formal contracts in order to increase predictability in interactions or expectations within exchange relationships [2].

Within the higher-order construct of relational governance, relational assurance mechanisms (RAMs), such as monitoring or reputation, are particularly known to increase predictability in interactions or expectations within (potential) exchange relationships [3-5]. According to Yamagishi and Yamagishi [6], assurance is defined as an expectation of benign behavior for reasons other than goodwill of the partner [7]. Hence, RAMs may be conceptualized as an important element of relational governance [3-5] although evidence evolving from research is lacking.

We discovered two key gaps in assurance research. Firstly, investigations related to assurance are fragmented and largely independent of RAMs and assurance as a concept. These investigations do, however, offer insights on the relationship between the antecedents and effects of RAMs. Secondly, our data shows that RAMs lack a

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Lang, M.; Wiesche, M.; Krcmar, H. (2017): Conceptualization of Relational Assurance Mechanisms - A Literature Review on Relational Assurance Mechanisms, Their Antecedents and Effects, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 852-866

conceptual consensus. Research is at odds when it comes to a consistent interpretation of the effects of RAMs. It is difficult to advance the theoretical and empirical investigation of RAMs, as existing literature does not provide a coherent and cumulative body of work. The gaps we discovered need to be considered when investigating RAMs as an important element of relational governance. In order to address these gaps, this article attempts to answer the following research questions (RQ). *RQ1: What mechanisms of assurance are exemplary discussed in information systems literature? RQ2: Which concepts are relevant when investigating assurance mechanisms and how are these concepts related?* To reach answers to these two questions, we conducted a systematic literature review and analyzed the results of this review in a structured manner.

Using our analysis results, we provide an overview of and conceptualize RAMs as published in IS literature. Furthermore, we point out identified concerns as the antecedents of RAMs, and the effects of RAMs on individuals within a theoretical framework.

The remainder of this article is structured as follows: In the next section, we describe the design of our literature review, including our methods for selecting journals and articles, and the subsequent analysis of the selected articles. Next, we discuss the theoretical background of our work including a psychological perspective of control as a source of assurance, and subsequently present the findings of our literature review. In the final section of the paper, we discuss our findings, address their theoretical implications and identify the limitations of this study.

2 Methodology

To identify relevant literature regarding our RQ1 and RQ2, we conducted a systematic literature review following the guidelines of Vom Brocke, Simons, Niehaves, Riemer, Plattfaut and Cleven [8] for the literature search, Webster and Watson [9] for literature analysis and synthesis, and Müller-Bloch and Kranz [10] to identify the research gap. According to our RQ1, the primary focus of this review is IS literature, identifying the key-concepts regarding our RQs within this research domain. Hence, the initial set of possible journals was limited to IS journals. As a result, all journals of the AIS senior scholars' "basket of 8 journals" were selected. To consider upcoming research topics as well, we also included high-quality, relevant articles from IS conferences.

We scanned journals using the online literature database EBSCOhost, searching for the term *"assurance"* used in the title, abstract, or keywords. For IS conference proceedings, we used the databases AISELNET and IEEE Xplore and searched abstracts for the word *"assurance"*. Articles published before June 2016 were considered. In order to get a broad overview of the concept "assurance" within exchange relationships, the search string was not limited further. As described below, further restrictions were carried out manually as part of the check for topic relevance. Overall, we initially identified 185 articles.

The articles were screened for relevance by reading title, abstract and, if necessary, the full text. In terms of our research, article relevance was defined as: the article uses the construct “assurance” in an exchange relationship context. Therefore, our selection comprises full research articles focusing on inter-organizational relationships, relationships between organizations and people, and inter-personal relationships. We excluded articles focusing on software development or product quality assurance as those do not cover assurance within an exchange relationship context. As a result, a set of 36 articles were included in our analysis. Next, we applied backward and forward search techniques to identify additional articles relevant for our research [8]. In the backward search, we reviewed the reference lists in our set of articles for appropriate articles. Similarly, we reviewed the citations of the articles in our set in Google Scholar. This final search technique yielded a final set of 52 articles.

After having identified the set of relevant articles, two researchers independently reviewed each article and developed an appropriate coding scheme. The researchers then compared their results and discussed any differences in their findings [9]. After three iterations, the researchers agreed on a final coding scheme, which was used for our analysis. This scheme included the used RAM, concerns as RAM antecedents (privacy concerns, security concerns, business integrity concerns), and the effects of the RAM on individuals (beliefs, intentions, behaviors) [10]. According RQ1 and RQ2, this research addresses a “knowledge void” research gap [10]. The final coding is summarized in a table (see Table 3 in the Appendix).

3 Theoretical Background

3.1 Assurance about Partners’ Intentions

Assurance is defined “as an expectation of benign behavior for reasons other than goodwill of the partner” [6]. Therefore, assurance is based on the knowledge of the incentive structure surrounding the relationship of two parties [6]. Such knowledge is particularly important in situations with high environmental uncertainty in which an actor does not have the capability of correctly detecting the partner’s intentions [11].

To gain knowledge of the incentive structure surrounding a (potential) relationship, individuals seek sources which provide additional information about (potential) partners [12]. These sources either accumulate information sufficient for allowing to be certain about (potential) partner’s intentions, provide deterrence against unilateral defection, or induce the partner to take a certain course of action with the use of strategies such as “tit-for-tat” [6, 13, 14]. Each source increase predictability in interactions or expectations within (potential) exchange relationships for reasons other than only the goodwill of the partner.

3.2 A Psychological Perspective of Control as a Source of Assurance

Research on assurance which considers the knowledge about the incentive structure surrounding (potential) relationships is based on a control agency perspective. In particular, this perspective allows not only an examination of the effects of personal control in which the individual acts as an assurance agent to protect information, but also includes proxy control and collective control [15, 16]. In proxy control, powerful others (such as the government and industry regulators) act as the assurance agents [15, 16]. In collective control, a collective acts as the assurance agent [16].

The personal control approach aims to directly assure outcomes from a client's perspective. People experience greater autonomy when they exercise direct personal control as the assurance agent [15-17]. Such control empowers individuals with mutual control over how their data and information, for example, may be used by service providers via technological and non-technological self-protection approaches [6, 15]. By using personal control, actors induce the partner to take a certain course of action with the use of strategies such as "tit-for-tat" [13, 18, 19]. Using these strategies, actors match their own behaviors to those displayed by personal control mechanisms (e.g. cooperating or trustful versus competing or opportunistic) [13].

The proxy control approach aims to indirectly assure outcomes via powerful others [20-22]. Institutional mechanisms are used from partners with few resources or low power to gain assurance through skillful and powerful third parties (e.g. industry self-control or legislation) [16, 23]. These mechanisms enable partners to access resources from third parties, such as knowledge and power, to assure outcomes. In case of opportunistic behavior, these assurance structure provide mechanisms of voice and recourse for the betrayed, which could create strong incentives for firms to refrain from opportunistic behavior and behave appropriately [14, 19, 24].

In the collective control approach, an individual, as a member of a group or collective that serve as an assurance agent, attempts to control the environment or outsiders. In collective control, responsibility, as well as agency, will be diffused among actors [25]. In the collective control approach, individuals attempt to share responsibilities among actors, internalize reference groups, and use their collective knowledge for decision making [16, 26]. Therefore, the collective is responsible for possible positive and negative outcomes to the same extent [16].

4 Findings

We adopted a psychological perspective of control and developed a theoretical framework for RAMs, its antecedents, and effects to provide a comprehensive overview and conceptual consensus for RAMs.

Therefore, the theoretical framework (Figure 1) posits that three sets of RAMs – personal control, proxy control, and collective control – influence individuals' beliefs, intentions, and behaviors when concerns are in place.

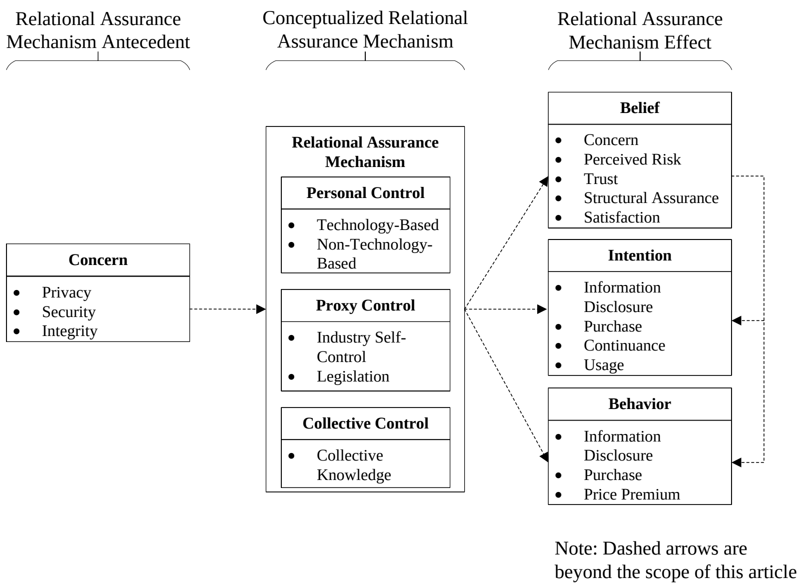


Figure 1. Theoretical Framework for Relational Assurance Mechanisms

Within the following sections, we outline the conceptualization of RAM, its antecedents, and effects in detail.

4.1 Conceptualization of Relational Assurance Mechanism

RAMs provide information about the incentive structure of (potential) partners and therefore, increase predictability in interactions or expectations within (potential) exchange relationships. According this notion, Table 1 summarizes the identified examples of RAMs using the key term "assurance" from our literature review. To distinguish the different examples of RAMs we provide a clear definition for each.

Table 1. Identified Relational Assurance Mechanism Examples and their Definitions

Example	**Definition**	**Source**
Certification	Defines an endorsement from a third-party organization attesting that a (potential) partner adheres to the organization's policy and a set of standards.	[27]
Corporative norm	Cooperative norms are defined as the values, standards, and principles to which a population of organizations adheres.	[28]
Feedback mechanism	Feedback mechanisms accumulate and disseminate information about the past trading behavior of organizations.	[28]
Law	Mandatory legal rules to ensure adequate protection of information.	[15]
Monitoring	A set of activities undertaken to assure that all transactions are performed as specified by a predetermined set of widely accepted agreements and rules.	[28]

Table 1. Identified Relational Assurance Mechanism Examples and their Definitions (Continued)

Example	Definition	Source
Persona-lization	Former mechanism which comprises tools and approaches that enable individuals to directly control outcomes.	[15]
Product description	The extent to which a consumer believes that a website is helpful in terms of fully evaluating a product.	[29]
Redundancy	The inclusion of extra components, which are not strictly necessary to functioning, in case of failure of other components.	[30]
Recommen-dation	A suggestion or proposal as to the best course of action.	[31]
Reputation	Reputation is imperfect and indirect information about a potential partner's traits.	[6]
Site quality	Reflects consumers' overall perceptions of how well they think a site works and looks, particularly in comparison to other sites.	[32]
Social Influence	Individual perceives support in decision making from his or her colleagues and others whose opinions matter.	[26]
Standardi-zation	The extent to which rules, procedures, and standards exist to guide the conduct of an activity and to evaluate performance.	[33]
Statement	A statement supplied by a (potential) partner that provides argumentation and claims to address certain concerns (e.g. privacy concerns).	[34]
Warranty	A warranty signals service quality and provides consumers some assurance in case of service failure.	[35]

Drawing on the work of Yamaguchi [16] on the differentiation of assurance agent perspectives, we conceptualize RAMs using the assurance agent perspectives personal control, proxy control, and collective control and highlight prominent paper examples.

Within personal control, individuals strive for primary control over their environment. For this assurance agent, literature suggest two major types of RAMs: technology-based and non-technology-based approaches [20]. Technology-based approaches include features such as monitoring, personalization, or technology redundancy (e.g. [17, 36]). Non-technological-based approaches are reading corporative norms, product descriptions or statements, providing direct feedback, considering existing warranties, site-quality, or standardization practices (e.g. [36]).

Proxy control describes institutional-based assurance of control whereby powerful forces act as the assurance agents. According to literature, individuals particularly rely on industry self-regulation and legislation to exercise proxy control [15]. Our research identified the use of specific certifications and laws as examples of industry self-regulation and legislation RAMs (e.g. [15]).

In collective control, one attempts to control the environment or outsiders as a member of a group or collective, which serves as an assurance agent. According to Yamaguchi [16], individuals "believe they are more efficacious as a collective than as an individual person". Therefore, individuals use their collective knowledge as a RAM to indirectly control the environment or outsiders. While reputation provides assurance for committed individuals to deal with uncertainty when involved with outsiders, social influence refers to an "individual's internalization of the reference group's subjective culture, and specific interpersonal agreements that the individual has made with others, in specific social situations" [6, 26] (e.g. [27]). Furthermore, by using the collective knowledge provided from internal or external sources, such as recommendations or reviews via feedback mechanisms, individuals overcome their concerns and adopt or continue a relationship [36, 37] (e.g. [38]).

Based on the assurance agent perspective, Table 2 summarizes our conceptualization of RAMs and identifies examples of these mechanisms from our literature review.

Table 2. Conceptualization of Relational Assurance Mechanisms

Assurance Agent	**Relational Assurance Mechanism**	**Identified Examples**
Personal Control	Technology-Based	Monitoring, Personalization, Redundancy
	Non-Technology Based	Corporative Norm, Product Description, Site-Quality, Feedback Mechanism, Standardization, Statement, Warranty
Proxy Control	Industry Self-Regulation	Certification
	Legislation	Law
Collective Control	Collective Knowledge	Reputation, Social Influence, Recommendation, Feedback Mechanism

In order to gain insights about how RAM concepts are interrelated, we next discuss the antecedents of RAMs as identified in literature.

4.2 Concerns as Antecedents of Relational Assurance Mechanisms

Based on the selected literature, we were able to identify three types of concerns that rise an individual's need for RAMs: privacy concerns, security concerns, and business integrity concerns. In the following section, we briefly explain each concern.

Privacy concerns are a primary concern dimension within IS literature, particularly in online transactions [15, 21, 36, 38, 39]. Privacy concerns within an online context are defined as individuals' concerns about the threat to their information privacy when submitting their personal information on the internet [36, 38]. Studies have identified that as privacy concerns increase, individuals seek RAMs [38, 40]; contrastingly, RAMs will lead to lower privacy concerns [15, 39]. Hence, privacy concerns and the presence of RAMs are highly negatively correlated.

Another antecedent of assurance identified in our review are *security concerns* [17, 36, 39, 41]. Based on the dimensions provided by Kim, Sivasailam and Rao [42], we distinguish between three types of security concerns: general security issues, transaction integrity, and authenticity of parties to transact. General security issues consist of insider abuse, unauthorized access, distributed denial of service attacks, and malware [17, 28, 36]. Transaction integrity is based on deletion, duplication, or alteration of documents [39, 43]. Alteration of documents refers to identity theft or authentication issues [44]. Security concerns depend not only on the security level of a firm, but also on the knowledge of individuals: e.g., how effective does the individual perceive the security protection mechanisms to be [39, 45].

Business integrity concerns are almost neglected within IS research even if such concerns have been identified as highly significant inhibitors for adoption decisions [39]. Such concerns are related to how (potential) partners (re-)use collected information from their customers and the possibility that a person or company may not fulfil a promise or complete a task. Especially within high environmental uncertainty, such concerns occur as a result of information asymmetry between (potential) exchange partners [36]. Such concerns may be amplified by the exponential proliferation of online scams and fake websites [42].

In the following section we outline the effects of RAMs on individuals as presented in our literature set.

4.3 Effects of Relational Assurance Mechanisms

This section outlines the effects of RAMs on an individual's beliefs (concern, perceived risk, trust, structural assurance, and satisfaction), intentions (information disclosure, purchase, continuance, and usage), and behaviors (information disclosure, purchase, price premiums).

First, RAMs affect an individual's *beliefs*. As discussed above, RAMs are in place to address certain concerns and therefore, researchers have also examined the effects of RAMs on concerns itself. RAMs, such as laws, certifications, and statements, have negative effects on an individual's concerns [15, 19, 21]. According to Xu, Dinev, Smith and Hart [19], concerns are partly mediated by the individual's perceived sense of control or perceived risk. Furthermore, related to concerns, studies identified the negative effect of product description, site quality, and certification on an indivudal's perceived uncertainty and perceived privacy risk [19, 29, 46]. Contrary to these negative effects, positive effects from RAMs, like certification or statements on trust, have been investigated [34, 36, 47]. Studies point out the positive effects of RAMs on structural assurance beliefs. Structural assurance is defined as the belief that success is likely because contextual conditions, such as statements, certifications and warranties, are in place [48]. Hence, structural assurance represents the perceived effectiveness of RAMs which are in place [49]. Lastly, researchers identified positive effects of perceived monitoring, perceived feedback, and cooperative norms on individual satisfaction with services or products [28].

Second, RAMs affect an individual's *intentions*. All of our identified studies on individuals' intentions considered trusting beliefs as mediators. Such studies point out the positive effects of RAMs, such as statements and site quality, on an individual's intention to disclose information [38, 50]. Furthermore, researchers identified positive effects of RAMs on purchase intentions [36, 39], intention to continue the relationship [28] or intention to use a web site [50]. Since, individuals tend to avoid losses, future research may consider control or risk perceptions as mediators to better explain an individual's intentions [51, 52].

Third, RAMs affect an individual's *behavior*. Studies identified the positive effects of privacy statements, certification, and customization on actual information disclosure [21, 36] and Oezpolat, Gao, Jank and Viswanathan [40] identified the positive effects of certifications on purchasing behavior. Dimoka, Hong and Pavlou [29] identified that product description and certification positivly influence the behavior to pay price premiums. Since the actual behavior can differ from an indiviudal's beliefs and intentions, further research is needed on how RAMs affect an indivudal's behavior [53].

5 Conclusion

This research was motivated by a fragmented body of knowledge, in which recent investigations largely examined assurance independently from the mechanisms and the concept itself. Based on this fragmented research, a conceptual consensus for RAMs is missing, even if RAMs are an important element of relational governance. To address these gaps, we conducted a systematic literature review, and identified examples of RAMs, as reported in IS literature. Based on this comprehensive overview, our subsequent analysis provides a conceptualization of RAMs. Last, our theoretical framework of RAMs further provides insights about antecedents and effects resulting from RAMs.

Before we conclude our major contributions, certain limitations should be considered when interpreting the results. Our literature review focused on RAMs as an important element of relational governance [3-5]. We recognize there are other forms of relational governance mechanisms such as joint actions or trust. While our theoretical arguments should extend to the instantiations of these other mechanisms of relational governance, more empirical work is needed to increase predictability in interactions or expectations within (potential) exchange relationships. Further investigations should particular build on the work of Yamagishi and Yamagishi [6], who distinguish between trust and assurance by taking social uncertainty into account. They claim, assurance is particular important in situations with low social uncertainty, while trust is needed when social uncertainty is high [6]. Another possible area of interest is to consider the influence of RAMs over time. Prior studies already found changes in the relevance of uncertainty for formal governance mechanisms [54, 55].

Our main contribution to the conceptualization framework of RAMs is threefold. First, we provide insights of the interrelation of existing assurance research and offer insights into how RAMs can be conceptualized. Second, we provide a theoretical framework to consider the concepts of RAMs and how these concepts are related to the antecedents and effects of RAMs. Third, we contribute to practice by providing an overview of existing RAMs and their effects [56]. Such findings might be used by practitioners, like security managers or auditing authorities, in order to adopt effective RAMs to increase predictability in interactions within exchange relationships.

6 Acknowledgements

This research was funded by the German Federal Ministry for Education and Research (grant. No. 16KIS0078).

Appendix

Table 3. Overview of Assurance Research

Citation	Relational Assurance Mechanism Examples [1]	Antecedent			Effect											
		Concern			Belief					Intention				Behavior		
		Privacy	Security	Business Integrity	Trust	Structural Assurance	Satisfaction	Concern	Risks	Information Disclosure	Purchase	Continuance	Usage	Information Disclosure	Purchase	Price Premium
[57]	SI	x	x			x	x						x			
[33]	Stand							x								
[58]	FM; Rep				x											x
[59]	Stat; SQ; Cert; CN; Rep	x			x					x						
[38]	Stat; SQ; Cert; CN; Rep; Rec	x			x					x						
[60]		x	x			x							x			
[61]		x	x		x	x							x			
[62]	SI	x	x	x		x		x								
[63]			x			x	x									
[29]	PD; Cert; FM; W			x				x								x
[5]	M															
[64]	Rep; SQ	x	x		x	x			x	x	x					
[21]	Stat; Cert	x						x						x		
[17]	Rec; Pers; SI		x													
[37]	Cert; Rec; SQ; SI	x							x		x		x			
[65]	Cert; SQ	x			x						x		x			
[36]	Cert; Rec; SQ; Pers; FM; Stat; SI	x	x		x	x								x		
[44]	Cert	x	x	x												
[47]	Stat	x	x		x											
[34]	Stat	x	x		x											
[39]	Cert; W	x	x	x							x					
[66]	Cert	x	x	x												
[67]	Stat	x														
[68]	L	x												x		
[69]	FM; Rep				x	x					x					
[70]	SQ; Rep	x	x		x	x										

[71]	Cert; FM; PD; W	x	x								x				
[72]		x	x		x										
[32]	SQ; Rep	x	x		x										
[73]		x	x		x	x									
[27]	SQ; Rep	x	x		x	x			x	x	x				
[50]	SQ; Rep	x	x		x	x				x	x				
[49]	Stat; Pers	x				x		x	x					x	
[40]	Cert	x	x	x											x
[28]	M; FM; CN		x				x		x		x	x			
[46]	PD; SQ; Rep; W; SI	x	x	x				x			x				x
[74]	Cert	x	x	x	x	x			x		x			x	
[20]		x													
[75]		x							x				x		
[76]	Cert		x	x											
[43]	Cert		x	x											
[41]	Cert		x												
[77]	Cert	x	x	x											
[18]	Red		x												
[19]	Stat; Cert	x						x	x						
[78]	Stat; Cert; L; Pers; Rep	x						x					x		
[15]	Stat; Cert; L; Pers; Rep	x						x							
[79]	SI		x			x	x								

[1] Cert = Certification, CN = Corporative norm, FM = Feedback mechanism, L = Law, M = Monitoring, Pers = Personalization, PD = Product description, Red = Redundancy, Rec = Recommendation, Rep = Reputation, SQ = Site quality, SI = Social Influence, Stand = Standardization, Stat = Statement, W = Warranty

References

1. Gopal, A., Koka, B.R.: The asymmetric benefits of relational flexibility: Evidence from software development outsourcing. Management Information Systems Quarterly 36, 553-576 (2012)
2. Poppo, L., Zenger, T.: Do formal contracts and relational governance function as substitutes or complements? Strategic Management Journal 23, 707-725 (2002)
3. Noordewier, T.G., John, G., Nevin, J.R.: Performance outcomes of purchasing arrangements in industrial buyer-vendor relationships. Journal of Marketing 54, 80-93 (1990)
4. Dyer, J.H.: Effective interfirm collaboration: How firms minimize transaction costs and maximize transaction value. Strategic Management Journal 18, 535-556 (1997)
5. Gundlach, G.T., Cannon, J.P.: “Trust but verify”? The performance implications of verification strategies in trusting relationships. Journal of the Academy of Marketing Science 38, 399-417 (2010)
6. Yamagishi, T., Yamagishi, M.: Trust and commitment in the United States and Japan. Motivation and Emotion 18, 129-166 (1994)
7. Barber, B.: The logic and limits of trust. Rutgers University Press, New Brunswick (1983)
8. Vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the giant: On the importance of rigour in documenting the literature search process. European Conference on Information Systems vol. 9, pp. 2206-2217, Verona (2009)

9. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. Management Information Systems Quarterly 26, 3 (2002)
10. Müller-Bloch, C., Kranz, J.: A framework for rigorously identifying research gaps in qualitative literature reviews. International Conference on Information Systems, Fort Worth (2015)
11. Rindfleisch, A., Heide, J.B.: Transaction cost analysis: Past, present, and future applications. Journal of Marketing 61, 30-54 (1997)
12. Williams, T.: Interorganisational information systems: Issues affecting interorganisational cooperation. The Journal of Strategic Information Systems 6, 231-250 (1997)
13. Axelrod, R., Hamilton, W.D.: The evolution of cooperation. Science 211, 1390-1396 (1981)
14. Shapiro, D.L., Sheppard, B.H., Cheraskin, L.: Business on a handshake. Negotiation journal 8, 365-377 (1992)
15. Xu, H., Teo, H.-H., Tan, B.C., Agarwal, R.: Research note - Effects of individual self-protection, industry self-regulation, and government regulation on privacy concerns: a study of location-based services. Information Systems Research 23, 1342-1363 (2012)
16. Yamaguchi, S.: Culture and control orientations. Oxford University Press, New York (2001)
17. Johnston, A.C., Warkentin, M.: Fear appeals and information security behaviors: an empirical study. Management Information Systems Quarterly 34, 549-566 (2010)
18. Wang, J., Chaudhury, A., Rao, H.R.: Research note - A value-at-risk approach to information security investment. Information Systems Research 19, 106-120 (2008)
19. Xu, H., Dinev, T., Smith, J., Hart, P.: Information privacy concerns: Linking individual perceptions with institutional privacy assurances. Journal of the Association for Information Systems 12, 798-824 (2011)
20. Son, J.-Y., Kim, S.S.: Internet users' information privacy-protective responses: A taxonomy and a nomological model. Management Information Systems Quarterly 32, 503-529 (2008)
21. Hui, K.-L., Teo, H.H., Lee, S.-Y.T.: The value of privacy assurance: An exploratory field experiment. Management Information Systems Quarterly 31, 19-33 (2007)
22. Tang, Z., Hu, Y., Smith, M.D.: Gaining trust through online privacy protection: Self-regulation, mandatory standards, or caveat emptor. Journal of Management Information Systems 24, 153-173 (2008)
23. Bandura, A.: Social cognitive theory: An agentic perspective. Annual Review of Psychology 52, 1-26 (2001)
24. Benassi, P.: TRUSTe: an online privacy seal program. Communications of the ACM 42, 56-59 (1999)
25. Latané, B., Darley, J.M.: The unresponsive bystander: Why doesn't he help? Prentice Hall (1970)
26. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: Toward a unified view. Management Information Systems Quarterly 425-478 (2003)
27. McKnight, D.H., Choudhury, V., Kacmar, C.: Developing and validating trust measures for e-commerce: An integrative typology. Information Systems Research 13, 334-359 (2002)
28. Pavlou, P.A.: Institution-based trust in interorganizational exchange relationships: The role of online B2B marketplaces on trust formation. The Journal of Strategic Information Systems 11, 215-243 (2002)

29. Dimoka, A., Hong, Y., Pavlou, P.A.: On product uncertainty in online markets: Theory and evidence. Management Information Systems Quarterly 36, (2012)
30. Burt, R.S.: Structural holes: The social structure of competition. Harvard University Press (2009)
31. Xiao, B., Benbasat, I.: E-commerce product recommendation agents: Use, characteristics, and impact. Management Information Systems Quarterly 31, 137-209 (2007)
32. Lowry, P.B., Vance, A., Moody, G., Beckman, B., Read, A.: Explaining and predicting the impact of branding alliances and web site quality on initial consumer trust of e-commerce web sites. Journal of Management Information Systems 24, 199-224 (2008)
33. Aubert, B.A., Houde, J.-F., Patry, M., Rivard, S.: A multi-level investigation of information technology outsourcing. The Journal of Strategic Information Systems 21, 233-244 (2012)
34. Kim, D., Benbasat, I.: Trust-assuring arguments in B2C e-commerce: Impact of content, source, and price on trust. Journal of Management Information Systems 26, 175-206 (2009)
35. Purohit, D., Srivastava, J.: Effect of manufacturer reputation, retailer reputation, and product warranty on consumer judgments of product quality: A cue diagnosticity framework. Journal of Consumer Psychology 10, 123-134 (2001)
36. Keith, M.J., Babb, J.S., Lowry, P.B., Furner, C.P., Abdullat, A.: The role of mobile-computing self-efficacy in consumer information disclosure. Information Systems Journal 25, 637-667 (2015)
37. Keith, M.J., Babb Jr, J.S., Furner, C.P., Abdullat, A.: Privacy assurance and network effects in the adoption of location-based services: An iPhone experiment. International Conference on Information Systems, pp. 237, St. Louis (2010)
38. Bansal, G., Zahedi, F., Gefen, D.: The role of privacy assurance mechanisms in building trust and the moderating role of privacy concern. European Journal of Information Systems 24, 624-644 (2015)
39. Kim, D.J., Yim, M.-S., Sugumaran, V., Rao, H.R.: Web assurance seal services, trust and consumers' concerns: An investigation of e-commerce transaction intentions across two nations. European Journal of Information Systems 1, (2015)
40. Oezpolat, K., Gao, G., Jank, W., Viswanathan, S.: The value of third-party assurance seals in online retailing: An empirical investigation. Information Systems Research 24, 1100-1111 (2013)
41. Sun, L., Srivastava, R.P., Mock, T.J.: An information systems security risk assessment model under the Dempster-Shafer theory of belief functions. Journal of Management Information Systems 22, 109-142 (2006)
42. Kim, D.J., Sivasailam, N., Rao, H.R.: Information assurance in B2C websites for information goods/services. Electronic Markets 14, 344-359 (2004)
43. Srivastava, R.P., Mock, T.J.: Evidential reasoning for WebTrust assurance services. Journal of Management Information Systems 16, 11-32 (1999)
44. Khazanchi, D., Sutton, S.G.: Assurance services for business-to-business electronic commerce: a framework and implications. Journal of the Association for Information Systems 1, 11 (2001)
45. Kim, D.J.: Self-perception-based versus transference-based trust determinants in computer-mediated transactions: A cross-cultural comparison study. Journal of Management Information Systems 24, 13-45 (2008)

46. Pavlou, P.A., Liang, H., Xue, Y.: Understanding and mitigating uncertainty in online environments: A principal-agent perspective. Management Information Systems Quarterly 31, 105-136 (2006)
47. Kim, D., Benbasat, I.: The effects of trust-assuring arguments on consumer trust in Internet stores: Application of Toulmin's model of argumentation. Information Systems Research 17, 286-300 (2006)
48. McKnight, D.H., Cummings, L.L., Chervany, N.L.: Initial trust formation in new organizational relationships. Academy of Management Review 23, 473-490 (1998)
49. Mousavizadeh, M., Kim, D.: A study of the effect of privacy assurance mechanisms on self-disclosure in social networking sites from the view of protection motivation theory. International Conference on Information Systems Fort Worth (2015)
50. McKnight, D.H., Choudhury, V., Kacmar, C.: The impact of initial consumer trust on intentions to transact with a web site: a trust building model. The Journal of Strategic Information Systems 11, 297-323 (2002)
51. Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk. Econometrica 47, 263-291 (1979)
52. Wiesche, M., Schermann, M., Krcmar, H.: Understanding the enabling design of IT risk management processes. International Conference on Information Systems, Fort Worth (2015)
53. Sheeran, P.: Intention - behavior relations: A conceptual and empirical review. European review of social psychology 12, 1-36 (2002)
54. Pflügler, C., Wiesche, M., Krcmar, H.: Are we already in a mature ITO market? A longitudinal study on the effects of market maturity on ITO vendor project performance. International Conference on Information Systems, Fort Worth (2015)
55. Schermann, M., Dongus, K., Yetton, P., Krcmar, H.: The role of transaction cost economics in information technology outsourcing research: a meta-analysis of the choice of contract type. The Journal of Strategic Information Systems 25, 32-48 (2016)
56. Lang, M., Wiesche, M., Krcmar, H.: What are the most important criteria for cloud service provider selection? A Delphi study. European Conference on Information Systems, Istanbul (2016)
57. Akter, S., D'Ambra, J., Ray, P.: User perceived service quality of m-health services in developing countries. European Conference on Information Systems, Pretoria (2010)
58. Ba, S., Pavlou, P.A.: Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. Management Information Systems Quarterly 26, 243-268 (2002)
59. Bansal, G., Zahedi, F.: The moderating influence of privacy concern on the efficacy of privacy assurance mechanisms for building trust: A multiple-context investigation. International Conference on Information Systems, pp. 7, Paris (2008)
60. Chandra, S., Theng, Y.L., Lwin, M.O., Foo, S.S.-B.: Understanding collaborations in virtual world. Pacific Asia Conference on Information Systems, pp. 96, Taipei (2010)
61. Chen, Y.-H., Chien, S.-H.: Investigating factors influencing the use of e-government service. Americas Conference on Information Systems, pp. 695, San Francisco (2009)
62. Chen, C.C., Mitchell, A.: Improving the trust of users on social networking sites via self-construal traits. Americas Conference on Information Systems, pp. 5, Lima (2010)
63. Devaraj, S., Fan, M., Kohli, R.: Antecedents of B2C channel satisfaction and preference: validating e-commerce metrics. Information systems research 13, 316-333 (2002)

64. Huang, L.-T., Farn, C.-K., Yin, K.-L.: On initial trust building for ecommerce: Revisiting from the perspective of signal theory and trust transference. European Conference on Information Systems pp. 94, Regensburg (2005)
65. Keith, M.J., Babb Jr, J.S., Furner, C.P., Abdullat, A.: The role of mobile self-efficacy in the adoption of location-based applications: An iPhone experiment. Hawaii International Conference on System Sciences, pp. 1-10. IEEE, Manoa (2011)
66. Kimery, K.M., McCord, M.: Third-party assurances: The road to trust in online retailing. Hawaii International Conference on System Sciences, pp. 10 pp. IEEE, Big Island (2002)
67. Kim, D., Koohikamali, M.: Does information sensitivity make a difference? Mobile applications' privacy statements: A text mining approach. Americas Conference on Information Systems, Savannah (2015)
68. Krasnova, H., Veltri, N.F.: Privacy calculus on social networking sites: Explorative evidence from Germany and USA. Hawaii international conference on System Sciences pp. 1-10. IEEE, Koloa (2010)
69. Kuan, H.-H., Bock, G.-W.: An exploratory study of before-interaction trust transference in multichannel retailers. International Conference on Information Systems, Las Vegas (2005)
70. Liao, Z.: Trust building and sustainable internet banking. Americas Conference on Information Systems pp. 16, Omaha (2005)
71. Li, E.Y., Yen, H.R., Liu, C.-C., Chang, L.F.: From structural assurances to trusting beliefs: Validating persuasion principles in the context of online shopping. Pacific Asia Conference on Information Systems, pp. 127, Jeju Island (2013)
72. Lowry, P.B., Posey, C., Bennett, R.B.J., Roberts, T.L.: Leveraging fairness and reactance theories to deter reactive computer abuse following enhanced organisational information security policies: An empirical study of the influence of counterfactual reasoning and organisational trust. Information Systems Journal 25, 193-273 (2015)
73. Mäntymäki, M.: Exploring customers' post-adoption perceptions: A study on trust, commitment and related constructs in B2C online service context. Pacific Asia Conference on Information Systems, pp. 216, Suzhou (2008)
74. Salehan, M., Kim, D.J., Lee, J.-N.: Antecedents, processes and consequences of web assurance seals: A meta-analysis approach. Pacific Asia Conference on Information Systems, Singapore (2015)
75. Spears, J.L.: The effects of notice versus awareness: An empirical examination of an online consumer's privacy risk treatment. Hawaii International Conference on System Sciences, pp. 3229-3238. IEEE, Wailea (2013)
76. Srivastava, R.P., Mock, T.J.: Evidential reasoning for WebTrust assurance services. Hawaii International Conference on Systems Sciences, vol. Track5, pp. 10 pp., Maui (1999)
77. Sutton, S.G., Khazanchi, D., Hampton, C., Arnold, V.: Risk analysis in extended enterprise environments: Identification of critical risk factors in B2B e-commerce relationships. Journal of the Association for Information Systems 9, 151-174 (2008)
78. Xu, H., Teo, H.-H.: Alleviating consumers' privacy concerns in location-based services: A psychological control perspective. International Conference on Information Systems, pp. 64, Charlottesville (2004)
79. Yan, A., Solomon, S., Mirchandani, D., Lacity, M., Porra, J.: The role of service agent, service quality, and user satisfaction in self-service technology. International Conference on Information Systems, Milan (2013)

Wie IT-Security Matchplays als Awarenessmaßnahme die IT-Sicherheit verbessern können

Andreas Rieb[1], Marko Hofmann[1], Alexander Laux[1], Steffi Rudel[1]
und Ulrike Lechner[1]

[1] Universität der Bundeswehr München,
Professur für Wirtschaftsinformatik, München, Deutschland
{andreas.rieb,marko.hofmann,alexander.laux,steffi.rudel,
ulrike.lechner}@unibw.de

Abstract. „Operation Digitales Chamäleon" ist eine IT-Security Schulung in Form eines Serious Games. Zielgruppe des Spiels sind IT-Sicherheitsprofessionals. Teams entwerfen Angriffs- und Verteidigungsstrategien – eingebettet in einen Prozess von Schulung und Debriefing. Die vorliegende Arbeit adressiert die Frage, wie „Operation Digitales Chamäleon" die IT-Security Awareness bei IT-Sicherheitsprofessionals beeinflusst. Hierzu wird im ersten Teil das Design vorgestellt, welches im zweiten Teil um ausgewählte Ergebnisse der Evaluation von sieben Spielen zu Spielerlebnis, Wahrnehmung, Wissensgewinn und geplanten Verhaltensveränderungen ergänzt wird. „Operation Digitales Chamäleon" ist ein Format der IT-Security Matchplays, die im Rahmen des Forschungsprojekts VeSiKi entwickelt und validiert werden.

Keywords: IT-Sicherheit, Sensibilisierung, IT-Security Awareness, IT-Security Matchplay, Serious Game

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Rieb, A.; Hofmann, M.; Laux, A.; Rudel, S.; Lechner, U. (2017): Wie IT-Security Matchplays als Awarenessmaßnahme die IT-Sicherheit verbessern können, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 867-881

1 Einführung

Smartphone-Sticker, Kaffeetassen, Poster, Schulungen und Trainings – die Liste der Maßnahmen, die Anwender oder Mitarbeiter für das Thema IT-Sicherheit sensibilisieren sollen, ist lang. Spaß macht das Thema IT-Sicherheit meistens nicht. IT-Sicherheit Kritischer Infrastrukturen (KRITIS) ist ein neues Thema in der IT-Sicherheit, das die Absicherung von Produktionsanlagen, Logistikketten o.a. thematisiert. Stuxnet hat das Thema IT-Sicherheit für Kritische Infrastrukturen bekannt gemacht. So ist es notwendig, dass sich Mitarbeiter mit dem Thema IT-Sicherheit auseinandersetzen, die bisher die IT lediglich als Enabler im Rahmen der Unternehmensprozesse genutzt haben. Ebenso müssen sich IT-Professionals über IT-Sicherheit von Anlagen Gedanken machen, die bisher nicht als IT-Sicherheitsrisiko betrachtet wurden. Im Spannungsfeld von unpopulärer IT-Sicherheit und drängender Notwendigkeit Kritische Infrastrukturen abzusichern, soll „Operation Digitales Chamäleon“ Mitarbeiter sensibilisieren und befähigen, adäquat auf Bedrohungen der IT-Sicherheit zu reagieren.

„Operation Digitales Chamäleon“ ist eine IT-Sicherheitsschulung, die als ein Serious Game konzipiert und im Rahmen des Forschungsprojektes „Vernetzte IT-Sicherheit Kritischer Infrastrukturen” (VeSiKi) als Teil der IT-Security Matchplay Serie entwickelt und validiert wird. Das Serious Game basiert auf dem Format Wargaming und integriert Elemente der IT-Risiko- und IT-Bedrohungsanalyse. Ziel ist es, IT-Sicherheitsverantwortliche und IT-Sicherheitsprofessionals im Umgang mit Advanced Persistent Threats (APTs) zu schulen, wie sie typisch sind für KRITIS. APTs sind ausgeklügelte, dauerhafte und verschleierte Cyber-Bedrohungen [1]. Symantec beschreibt dies als “An advanced persistent threat (APT) uses multiple phases to break into a network, avoid detection, and harvest valuable information over the long term.“ und vermutet dass „Betrug – oder Schlimmeres“ Teil von APTs sind [2][3]. Zudem modifizieren die Angreifer während ihrer APT-Attacken regelmäßig ihre Angriffsvektoren und ihren Schadcode, um eine Entdeckung zu erschweren [4].

Dieser Artikel erweitert erste Publikationen zum Serious Game „Operation Digitales Chamäleon“. In [5] wurde „Operation Digitales Chamäleon“ als Instrument der Open Innovation mit einer Analyse des Innovationsgrads der APTs aus Spielergebnissen präsentiert. Das Spiel mit ausgewählten Spielergebnissen sind in einem Short Paper dargestellt [6].

2 State of the Art – IT-Security Awarenessmaßnahmen

„Operation Digitales Chamäleon“ wurde als IT-Security Awarenessmaßnahme konzipert. Im ersten Abschnitt wird der State of the Art in IT-Sicherheitsmaßnahmen vor allem in der Praxis zusammengefasst, während im zweiten Abschnitt vor allem der Stand der wissenschaftlichen Literatur zu Awareness dargestellt ist. In beiden Abschnitten sind die Verbindung von (passivem) „Wissen” zu (aktivem) „richtigen Handeln” sowie die Evaluation des Erfolgs besonders berücksichtigt. Der dritte Abschnitt ist „Serious Games” als Methode im Themenfeld IT-Sicherheit gewidmet.

2.1 IT-Security Awareness und Sensibilisierung in der Praxis

Maßnahmen zur Sensibilisierung für das Thema IT-Sicherheit (auch bezeichnet als IT-Security Awareness) sind in der Praxis weit verbreitet. So werden international bspw. in [7] und [8] verschiedene Maßnahmen inkl. Vor- und Nachteilen beschrieben. In einer Studie aus dem Jahr 2007 wurde u.a. untersucht, welche Maßnahmen in der Praxis eingesetzt und wie der Erfolg dieser Maßnahmen gemessen wird [9]. Das Bundesamt für Sicherheit in der Informationstechnik (BSI) bzw. die Allianz für Cyber-Sicherheit zu IT-Security Awareness stellt in ihrer Studie von 2015 [10] fest, dass 63% der Befragten angeben, dass IT-Security Awarenessmaßnahmen in der Organisation durchgeführt werden und dass Awarenessmaßnahmen überwiegend eine Präventivmaßnahme aufgrund wachsender Risiken sind. Als Maßnahmen werden angeführt (in absteigender Reihenfolge der Nennung): Dienstanweisungen/Vorschriften/Policies, Schulungen/Seminare in Gruppen, Informationskampagnen (z.B. Flyer, Poster), Mitarbeiterveranstaltungen (z.B. Roadshow), Einzelunterricht/E-Learning, Testszenarien zur Prüfung des Mitarbeiterverhaltens und andere; wobei als wesentliche Medien Intranet, E-Mail, Folienpräsentationen und Broschüren/Flyer, Poster und Mitarbeiterzeitschriften sowie Videos genannt werden. (Gewinn-)Spiele spielen nur eine untergeordnete Rolle. Diese Maßnahmen werden überwiegend nur sporadisch durchgeführt und eine Erfolgsmessung bleibt meistens aus. Nur ein kleiner Prozentsatz der befragten Unternehmen wertet die Sicherheitsvorfälle zur Erfolgsmessung der Awarenessmaßnahmen aus. Diese Zahlen und Daten decken sich mit der Erfahrung der Autoren dieses Artikels – interaktive Formate für IT-Security Awarenessmaßnahmen werden nur selten verwendet und der Nutzen für die IT-Sicherheit kaum evaluiert.

Für die IT-Sicherheit Kritischer Infrastrukturen stellt das BSI als maßgebliche staatliche Organisation in Deutschland Referenzwerke der Informationssicherheit bereit. Je nach Quelle ist auch von „IT-Sicherheit" die Rede, jedoch wird derzeit dieser Begriff durch den Begriff „Informationssicherheit" ersetzt. Grund dafür ist, dass sich IT-Sicherheit primär mit dem Schutz elektronisch gespeicherter Informationen und deren Verarbeitung beschäftigt, während Informationssicherheit als umfassender zu betrachten ist [11]. Innerhalb der o.a. Referenzwerke beschreibt das BSI auch Maßnahmen zur Sensibilisierung für das Thema Informationssicherheit und motiviert wie folgt: „Aufgrund der Entwicklung, dass inzwischen weder technische noch organisatorische Schutzmaßnahmen allein wirkungsvoll gegen Cyber-Bedrohungen schützen, rückt der Nutzer zunehmend in den Fokus. Er ist bei vielen Arten von Cyber-Angriffen noch häufig als das schwächste Glied in der Angriffskette zu sehen. […] Die kontinuierliche Sensibilisierung von Mitarbeitern für die bestehenden Risiken ist daher unerlässlich." [12]. Als Ziel wird genannt, die „Informationssicherheit in unser tägliches Handeln zu überführen". Als Aufgaben werden bspw. angeführt: „Vermittlung der Ziele der Informationssicherheit", „Empowerment, d.h. Vermittlung (praktischer) Kompetenzen hinsichtlich der Umsetzung von Regelungen" oder „Positionierung von Informationssicherheit durch Kommunikation von Security-Themen,-Aufgaben, -Tools und -Protagonisten mit dem

Ziel, Bekanntheit und Akzeptanz zu steigern bzw. in der Unternehmenskultur als Teil der Sicherheitskultur zu etablieren." [12].

Im Baustein der IT-Grundschutzkataloge zu „Sensibilisierung und Schulung zur Informationssicherheit" (B 1.13) [13] wird motiviert „alle Mitarbeiter erkennen und akzeptieren, dass [die Informationssicherheit] ein bedeutender und notwendiger Faktor für den Erfolg der Institution ist und [die Mitarbeiter] bereit sind, Sicherheitsmaßnahmen wirkungsvoll zu unterstützen". Die Akzeptanz von Sicherheitsmaßnahmen zielt auf „langfristige Verhaltensänderungen" ab, besonders wenn Informationssicherheit mit Komfort- oder Funktionseinbußen verbunden ist [13]. Planspiele werden als eine Maßnahme (M 3.47; [14]) empfohlen, um die Sensibilisierung und Schulung zur Informationssicherheit erfolgreich zu gestalten. Es ist wichtig, „eine positive und konstruktive Grundstimmung" in Planspielen zu haben – denn „Ständige Angst vor Sicherheitsvorfällen kann einerseits zur Verdrängung von Sicherheitsproblemen und andererseits zu Panikreaktionen verleiten." [14]. Das BSI sieht in seinen Referenzwerken der IT-Sicherheit Sensibilisierung der Mitarbeiter als wichtiges Element einer IT-Sicherheitsstrategie an und zeigt auf, dass Planspiele oder Rollenspiele ein geeignetes Instrument wären, Mitarbeiter in einer positiven Atmosphäre zu schulen – im Baustein Planspiele jedoch gibt es außer drei Themen wenig Konkretes zu Planspielen für IT-Sicherheit.

An dieser Stelle soll erwähnt werden, dass sowohl Planspiele als auch Rollenspiele Formen von Serious Games sind [15]. In der Praxis, wie auch in der Literatur werden diese Begrifflichkeiten jedoch häufig synonym verwendet [16].

2.2 IT-Security Awareness aus Sicht der Wissenschaft

Die ausgezeichnete Literaturübersicht von Hänsch und Benenson systematisiert die Definitionen der IT-Security Awareness in Wahrnehmung (Perception), Schutz (Protection) und Verhalten (Behavior) mit den Fähigkeiten Bedrohungen zu erkennen (recognize threat), Wissen über Lösungen zu haben (know solutions) und richtig zu handeln (act right) [17]. Betrachtet werden in dieser Übersicht auch Messkriterien für IT-Security Awareness. Während Wahrnehmung und Schutz – entsprechend dieser Literaturübersicht – in verschiedenen Ansätzen gemessen werden kann, finden sich zu IT-Security Awareness für „richtiges Handeln" nur wenige Methoden.

Die Forschung von Bulgurcu et al. illustriert die zentrale Rolle von IT-Security Awareness: Awareness beeinflusst mittelbar Nutzen von Compliance genau wie wahrgenommene Kosten von Compliance und von Non-Compliance [18]. Nach Johnston und Warketing genügen Wissen über Bedrohungen und Verwundbarkeiten nicht, sondern reduzieren Wirksamkeit und Selbstvertrauen (Response, Self Efficacy) und damit die Bereitschaft (richtig) zu handeln [19].

Das Konzept von IT-Security Literacy schlägt ähnlich wie IT-Security Awareness den Bogen zwischen „Wissen" und „Wahrnehmung" zu „richtigem Handeln" in der IT-Sicherheit speziell in Schulungen und Trainings (vgl. bspw. [20]).

2.3 Serious Games in der IT-Sicherheit

„Operation Digitales Chamäleon“ wurde unter anderem entwickelt, um IT-Sicherheitsverantwortliche und IT-Sicherheitsprofessionals in spielerischer Art und Weise zu sensibilisieren. Die Methode Wargaming, die wir für „Operation Digitales Chamäleon“ adaptieren, hat eine lange Tradition: In der Literatur hat Wargaming seinen Ursprung im 19. Jahrhundert, als Baron von Reisswitz diese Methode nutzte, um Kommandostäbe in ihrer Entscheidungsfindung ,,besonders in dynamischen und unvorhersehbaren Situationen zu schulen“ [21]. Charakteristisches Element des Wargamings ist die durch Regeln gesteuerte und von Schiedsrichtern bewertete Auseinandersetzung der Ideen und Pläne zumindest zweier Teams (Rot und Blau). Die Methode wird vor allem für effektives Training von Kommandostäben verwendet, seltener auch zum Erforschen komplett neuer Bedrohungen. Jedoch ist Wargaming in der Domäne IT-Sicherheit eher neu.

2012 gab ENISA einen Überblick über 85 IT-Security Übungen [22]. Die meisten davon waren theoretische Übungen “to validate plans and integration of procedures prior to moving on to more complex, team-based activities” [22]. Ziele dieser IT-Security Übungen waren u.a. die Steigerung von IT-Security Awareness bzgl. Cyber-Bedrohungen, Rollenidentifikation, Klärung von Verantwortlichkeiten und zuständigen Behörden, Entscheidungsfindung, Überprüfung des Incident Managements oder Vertrauensbildung zwischen Staaten [22]. 2015 wurde der ENISA-Bericht um mehr als 200 Übungen erweitert [23]. Hier hebt ENISA Methoden wie Red Teaming, Diskussionsbasierte Spiele, Capture the Flag, Seminare und weitere hervor. Gemäß ENISA [23] wird nur ein Bruchteil (11%) der Übungen mit gegenüberstehenden Teams durchgeführt.

Weitere Serious Games für IT-Sicherheit sind beispielsweise „Friend Inspector“ – ein softwarebasiertes Serious Game zur Steigerung der IT-Security Awareness unter Facebook-Nutzern [24][25]. Das Kartenspiel „Elevation of Privilege“ dient ebenfalls der Steigerung von IT-Security Awareness, jedoch stehen hier Softwareentwickler im Vordergrund [26]. Die Zielgruppe „Mitarbeiter eines Unternehmens“ adressiert das Kartenspiel von Beckers und Pape zu Gefahren und Methoden des Social Engineerings [27]. „Game of Threats“ adressiert das Management [28].

3 Methode

In der Forschung orientiert sich dieser Beitrag am Paradigma der gestaltungs-orientierten Forschung (Design Science) von Hevner et al. [29]. Wir verwenden einen iterativen Ansatz in Design und Evaluation mit einem kreativen Designprozess – so wie es bspw. Baskerville und Pries-Heje beschreiben [30]. Dieser kreative Designprozess bezieht insbesondere die mehr als 10 jährige Erfahrung des ersten Autors als IT-Security und IT-Security Awareness Spezialist mit ein.

In der Datengenerierung hat der erste Autor des vorliegenden Artikels als Spielleiter die Daten gesammelt, Umfragen durchgeführt und Beobachtungen protokolliert. In einem Fall wurde er von einem Beobachter in der Datensammlung unterstützt.

4 Das Spieldesign von „Operation Digitales Chamäleon”

„Operation Digitales Chamäleon“ ist als IT-Security Awarenessmaßnahme konzipiert und speziell entwickelt für die Schulung zum Thema IT-Sicherheit Kritischer Infrastrukturen. Diese Schulung umfasst neben der eigentlichen Spielphase eine Phase der Wissensvermittlung zu aktuellen IT-Sicherheitstechnologien und -bedrohungen mit ggf. Live-Hackings sowie eine Phase des Debriefings, die u.a. auch eine Umfrage beinhaltet. „Operation Digitales Chamäleon“ ist ein Planspiel und verwendet „traditionelle“, also nicht IT-gestützten Spielmaterialien.

4.1 Teams und Zielgruppe

„Operation Digitales Chamäleon“ ist als IT-Security Awarenessmaßnahme für IT-Sicherheitsverantwortliche und IT-Sicherheitsprofessionals als Zielgruppe konzipiert. Spielteilnehmer füllen in ihren Organisationen IT-Sicherheitsfunktionen aus oder haben berufliche Erfahrung im Themenfeld IT-Sicherheit (bspw. IT-Sicherheitsadministratoren auf operativer Ebene, IT-Security Manager). „Operation Digitales Chamäleon“ wird mit einer Gruppe von 8 bis 20 Personen durchgeführt.

Im Serious Game tritt diese Zielgruppe in Teams gegeneinander an: Team Rot als Angreifer gegen Team Blau als Verteidiger – ein weißes Team übernimmt Spielleitung und Schiedsrichterfunktion. Typischerweise haben Team Rot und Team Blau jeweils drei bis sieben Mitglieder.

Zu Beginn wählt Team Rot eine von fünf vorgegebenen Angreiferrollen (Threat Actors). Die Angreiferrolle beschreibt die Motivation und die Fähigkeiten des Threat Actors, die Grundlage für die zu planenden Angriffsstrategien bilden. In „Operation Digitales Chamäleon“ stehen insgesamt fünf verschiedene Angreiferrollen bzw. Threat Actors zur Verfügung: Script Kiddies, Cyber Criminals, Employees, Nation States und Hacktivists. Die Rolle „Hacktivists“ ist bspw. auf der Basis von Hald und Petersen [31] in den Spielunterlagen beschrieben: Ruhm, Ehre oder ein „moralischer Grund” sind die Motivationen, die finanziellen Ressourcen sind limitiert, jedoch sind die technischen Fähigkeiten „gut“.

Team Blau verteidigt die Kritische Infrastruktur und dem Team ist bekannt, welcher Threat Actor sie angreifen wird. So hat Team Blau die Aufgabe, sich in die Rolle des Angreifers hineinzuversetzen, die kritischen Assets (z.B. Personen, Netzwerkkomponenten, IT-Systeme u.a. Vermögenswerte) zu identifizieren und die IT-Sicherheitsmaßnahmen spezifisch auf zu erwartenden Angriffe auszurichten. Team Blau etabliert Schutzmaßnahmen zu den Faktoren Organisation, Technik und Mensch und erstellt ein IT-Sicherheitskonzept.

Team Weiß trifft als Spielleitung die Entscheidungen über den Spielablauf und überwacht die Einhaltung der Regeln. Beobachter oder Experten dokumentieren Spielverlauf und Ergebnisse und wirken bei der Ermittlung des Siegerteams mit.

Team Rot und Team Blau haben je einen Teamleiter, der durch Losverfahren bestimmt wird. Die Teamleitung organisiert die Teamarbeit, präsentiert die Ergebnisse und wählt (im Fall Rot) die Angreiferrolle.

4.2 Spielbrett und Spielmaterialien

„Operation Digitales Chamäleon“ verwendet konventionelle, nicht-digitale Spielmaterialien. Dazu gehören u.a. ein Spielbrett, auf dem ein Netzplan mit IT-Infrastrukturkomponenten abgebildet ist, wie sie für Kritische Infrastrukturen typisch sind (siehe Abbildung 1). Assets dieses Netzplans sind u.a. eine Industrieanlage sowie Windows PCs mit verschiedenen älteren, teilweise nicht mehr unterstützten Betriebssystemen. Zudem beinhaltet der Netzplan einen Fernwartungszugang, mobile Clients und ein Office-Teilnetz mit aktuelleren Betriebssystemen, Servern und Druckern. Die Teilnetze werden zu Beginn des Serious Games nur durch Router geschützt.

Abbildung 1: Das Spielbrett zu „Operation Digitales Chamäleon“

Mit Karten wird markiert, welches Asset der IT-Infrastruktur angegriffen bzw. verteidigt wird. Auf Post-Its werden Ideen zu Angriffsvektoren oder Schutzmaßnahmen notiert. Das Spielmaterial beinhaltet (pro Team) Karten, die die Rollen beschreiben, einen Satz Spielregeln, sowie Stifte für Notizen auf dem Spielbrett. Spielleibchen in den Farben Rot, Blau oder Weiß markieren die Teamzugehörigkeit. Dem Spielleiter bzw. dem weißen Team stehen für die Unterstützung der Entscheidung über die Machbarkeit diverser Angriffe oder Schutzmaßnahmen Zufallskarten mit unterschiedlichen Wahrscheinlichkeiten zur Verfügung.

4.3 Spielregeln, Spielablauf und Spielmissionen

Zu Beginn der Spielphase wird die Rahmenlage durch die Spielleitung präsentiert. Team Rot wird vorgegeben, dass der Angriff entsprechend der gewählten Spielerrolle eine Wirkung entfalten muss, die es notwendig macht, das BSI entsprechend den Meldepflichten – so wie sie im IT-Sicherheitsgesetz festgelegt sind – zu informieren [32].

Die Teams werden im Briefing ermuntert kreativ zu sein. Somit sind der Phantasie im Spiel nur wenig Grenzen gesetzt – es limitieren Spielzeit und Plausibilität.

Team Rot nimmt die Sicht des gewählten Threat Actors ein und soll einen Angriffspfad entwickeln, der für diese Angreiferrolle hinsichtlich Absichten, Ressourcen und Methoden typisch ist. Dieser Angriffspfad besteht aus voneinander abhängigen und ggf. alternativen Angriffsvektoren. In den weiteren Spielmissionen bekommt Team Rot die Aufgabe, Alternativen zu den einzelnen Angriffsvektoren zu definieren und den Angriffsbaum auch zu präsentieren.

Das blaue Team verfolgt das Ziel, die bevorstehenden Cyberangriffe erfolgreich abzuwehren. Team Blau entwickelt ein IT-Sicherheitskonzept sowie Schutzmaßnahmen zu den Faktoren Organisation, Technik und Mensch. Die Regeln des Spiels legen fest, dass die IT-Sicherheitsstrategie von Team Blau jedoch nicht zu Lasten der Verfügbarkeit und Nutzerfreundlichkeit der IT der KRITIS gehen darf. Die Regeln geben ebenfalls vor, dass die Mitarbeiter (der KRITIS) die Internetanbindung auch für Internetdienste wie Facebook nutzen. Ferner dürfen Netzverbindungen und Fernwartungszugänge nicht aus Gründen der IT-Sicherheit gekappt werden. Solche und andere der Realität entsprechenden Prämissen verhindern, dass Team Blau die „sicherste“ Lösung – das dauerhafte Trennen sämtlicher Internetverbindungen – wählt. Team Blau ist so gefordert, kreative Lösungen zu entwickeln, um die Angriffe von Team Rot zu antizipieren und abzuwehren.

In der Ermittlung des siegreichen Teams sind Nachvollziehbarkeit in der Argumentation und Transparenz der Entscheidungsfindung wichtig. Die Teams präsentieren die erarbeiteten Lösungen. Die Spielleitung geht Angriffsvektor um Angriffsvektor durch und bezieht die Teams in die Diskussion mit ein. In dieser Bewertung finden drei Kriterien Anwendung: (Technische) Machbarkeit (1), Plausibilität (2) und Erfolg im Hinblick auf Schutzmaßnahmen, die von Team Blau festgelegt wurden (3). Für Fälle, in denen keine eindeutige Entscheidung in der Diskussion getroffen werden kann, hilft „der Zufall“, implementiert als Spielkarten.

Die Teams erhalten einen Ordner mit Informationen über den Spielablauf und den Spielregeln. Diese legen beispielsweise fest, dass außerhalb der Spielzeiten nicht an der Weiterentwicklung der Strategie im Team gearbeitet werden darf. Zudem dürfen die Teams keine IT-Unterstützung in Form von Laptops, Smartphones o.ä. nutzen.

4.4 Das Debriefing

Ein Debriefing gibt den Teilnehmern die Möglichkeit, ihre persönlichen Erkenntnisse zu reflektieren und gibt Anregungen, das im Spiel Erfahrene in den Arbeitsalltag zu übernehmen. Hier folgt das Spiel den Anregungen von Kriz [33]. Wesentliche Ergebnisse, die im Rahmen dieser Debriefings gewonnen wurden, werden im nächsten Abschnitt vorgestellt.

5 Ergebnisse der „Operation Digitales Chamäleon“

Der nachfolgende Abschnitt stellt wesentliche Ergebnisse der gespielten Serious Games „Operation Digitales Chamäleon“ vor. Hierbei stützt sich die Empirie auf sieben Spiele, die in Tabelle 1 gelistet sind.

Tabelle 1. Durchgeführte Spiele „Operation Digitales Chamäleon"

#	Datum	Dauer Training	Dauer Spiel	Teil-nehmer	Anzahl rote Teams	Anzahl blaue Teams	Sektor	Länder-code
1	10/2015	3d	6h	11	1	1	Staat und Verwaltung (Polizei, Justiz)	DEU
2	01/2016	5d	9h	9	1	1	Transport und Verkehr (Luftfahrt)	FRA
3	03/2016	2d	6h	10	1	1	Staat und Verwaltung (Polizei)	DEU
4	03/2016	2d	10.5h	19	3	1	Staat und Verwaltung (Militär)	DEU
5	05/2016	2d	10h	18	2	1	Staat und Verwaltung (Militär)	DEU
6	05/2016	2d	10h	20	3	1	Staat und Verwaltung (Militär)	DEU
7	07/2016	2d	10.5h	17	3	1	Staat und Verwaltung (Militär)	DEU

Es ist zu beachten, dass in dem iterativen Vorgehen beim Design des Serious Games auch die Evaluation weiterentwickelt wird. So wurde der Fragebogen zur Evaluation des Spiels erst ab Spiel #4 eingesetzt. Ergebnisse wurden in den Spielen #2 und #3 mittels Post-Its im Rahmen einer moderierten Gruppendiskussion erhoben. Spiel #1 diente als Pretest mit Fokus auf die Spieldurchführung – hier wurden Notizen des Spielleiters während des Debriefings angefertigt.

5.1 Spielerlebnis

Spielerisches Lernen wird mit Spaß assoziiert und eine positive, angstfreie Atmosphäre ist entsprechend den Empfehlungen des BSI und den empirischen Resultaten (vgl. Kap. 2.1 und 2.2) wichtig für den Erfolg von Schulungen und speziell Sensibilisierungsmaßnahmen. Dies steht in Einklang mit den Empfehlungen zur Entwicklung von Serious Games (z.B. [34]). Gemäß McConigal ist Spaß während des Spielens eine wichtige Voraussetzung zur Steigerung der Motivation der Teilnehmer und der Qualität der Spielergebnisse [35]. Killmeyer definiert den Faktor Spaß ebenfalls als wichtige Voraussetzung für eine erfolgreiche IT-Sicherheitssensibilisierung [36]. Tabelle 2 zeigt die Ergebnisse der Spielevaluation per Fragebogen zur Aussage „Das Cyberwargame hat mir Spaß gemacht".

Tabelle 2. Spaßfaktor in Spiel #4 bis #7

Aussage	*Trifft nicht zu*	*Trifft eher nicht zu*	*Neutral*	*Trifft teilweise zu*	*Trifft voll zu*
Das Cyberwargame hat mir Spaß gemacht.	0	3	4	41	26

„Operation Digitales Chamäleon" hat den meisten Spielteilnehmern Spaß gemacht. Beobachtungen seitens des Spielleiters bestätigen dieses Ergebnis: Während der Durchführungen konnten des Öfteren Reaktionen wie „lautes Lachen" beobachtet werden. In allen Spielen herrschte eine konzentrierte, aber lockere Atmosphäre – keiner der 104 Spielteilnehmer der ersten sieben Spiele ist vorzeitig ausgestiegen. Eine positive Arbeitsstimmung ist wesentliche Voraussetzung für den (langfristigen) Erfolg der Sensibilisierungsmaßnahme und damit für die Verbesserung der IT-Sicherheit wie in den folgenden Kapiteln dargestellt.

5.2 Wahrnehmung

IT-Security Awareness beinhaltet die Wahrnehmung (Definitionen Awareness Kap. 2.2) von Gefahren und Risiken. Im Debriefing von Spiel #4 bis #7 wurde ermittelt, ob „Operation Digitales Chamäleon“ die Wahrnehmung verbessert.

Tabelle 3. Risikobewusstsein / Bedrohungseinschätzung in Spiel #4 bis #7

Aussage	*Trifft nicht zu*	*Trifft eher nicht zu*	*Neutral*	*Trifft teilweise zu*	*Trifft voll zu*
Ich stelle für mich eine Steigerung meines IT-Risikobewusstseins fest.	1	3	9	34	27
Ich kann die Komplexität von Bedrohungen auf die IT-Infrastruktur besser einschätzen.	0	4	9	46	15

Tabelle 3 stellt dar, dass „Operation Digitales Chamäleon“ sowohl das IT-Risikobewusstsein, als auch die Fähigkeit, die Komplexität von Bedrohungen auf die IT-Infrastruktur einschätzen zu können, positiv beeinflusst. Der Selbsteinschätzung der Teilnehmer nach, ist das Spiel nicht nur geeignet (Fakten-) Wissen zu vermitteln, sondern schlägt die Brücke zu Wahrnehmung und richtigem Handeln (vgl. Kap. 2.2).

5.3 Wissensgewinn

Im Debriefing von Spiel #4 bis #7 wird im Fragebogen nach dem Wissenszuwachs gefragt und die Mehrheit stellt einen Wissenszuwachs für sich fest (vgl. Tabelle 4).

Tabelle 4. Wissensgewinn in Spiel #4 bis #7

Aussage	*Trifft nicht zu*	*Trifft eher nicht zu*	*Neutral*	*Trifft teilweise zu*	*Trifft voll zu*
Ich stelle für mich einen Wissenszuwachs am Ende der Veranstaltung fest.	0	1	4	28	41

In einer zweiten, offenen Frage (Durchführung #4 bis #7) oder in einem Debriefing über Post-Its werden die Spielteilnehmer nach dem Wissensgewinn gefragt. Nach sieben Durchführungen stehen 144 Aussagen zur Auswertung zur Verfügung. In einer qualitativen Inhaltsanalyse nach Mayring [37] werden diese 144 Aussagen in die drei Hauptkategorien „Wissen über Angriffe“, „Wissen über Schutzmaßnahmen“ und „Sonstiges Wissen“ einsortiert.

Die beiden Kategorien „Wissen über Angriffe“ mit 57 Aussagen sowie „Wissen über Schutzmaßnahmen“ mit 66 Aussagen sind in etwa gleich stark ausgeprägt („Sonstiges Wissen“ mit 21 Aussagen). Dieses Ergebnis ist bemerkenswert: Die „böse Seite“ – also das rote Team – hat die kreativere Rolle und so wäre eine intensivere Auseinandersetzung mit Angriffen, also mehr Nennungen in der Kategorie „Wissen über Angriffe“ zu erwarten. Zudem sind in den sieben Spielen mehr rote als blaue Teams angetreten und zusätzlich werden in der Wissensvermittlung bspw. in Live-Hackings Angriffe plastischer behandelt als Schutzmaßnahmen. Dies illustriert, dass

„Operation Digitales Chamäleon“ zur Auseinandersetzung mit Schutzmaßnahmen anregt – entsprechend dem Forschungsziel, IT-Sicherheit für KRITIS zu verbessern.

Eine zweite Inhaltsanalyse der Aussagen zu „Technik“, „Organisation“, „Mensch“, und „Übergreifend“ lässt erkennen, dass die Spielteilnehmer auf Angreiferseite mehr „übergreifendes Wissen“ gewinnen konnten. Exemplarisch ist hier die Aussage „Arbeitsweise von Nation States“ zu nennen. Auf der Seite der Schutzmaßnahmen werden mehr spezifische Einzelmaßnahmen genannt. Genannte Beispiele sind das PAP-Prinzip des BSI (vgl. [38]) oder die Funktionsweise einer Datendiode (vgl. [39]).

Spielteilnehmer stellen interessanterweise nicht nur neueste Bedrohungen oder Schutztechnologien sondern auch ältere Angriffsvektoren oder etablierte Sicherheitstechnologien als Wissenszuwachs dar. So wurden als „Neues Wissen“ referernziert: der seit Jahrzehnten bekannte Angriffsvektor „Man-in-the-Middle“, der „Unterschied zwischen IDS / IPS“ , Angriffsvektoren wie „Man-in-the-Cloud“ [40] aus dem Jahr 2015 oder „Watering hole attack“ [41] von 2012. Das illustriert die Notwendigkeit für Sensibilisierungsmaßnahmen – auch für Funktionsträger.

„Operation Digitales Chamäleon“ vermag also neues Wissen zu vermitteln und sensibilisiert besonders im Themenfeld der Schutzmaßnahmen.

5.4 Verhaltensänderung

Im Debriefing reflektieren die Teilnehmer ihre Erfahrungen der Spielphase und leiten daraus individuell Vorsätze für ihre berufliche Tätigkeit ab. Die empirische Basis umfasst 208 Vorsätze (aus Spiel #1-#7), die mittels Fragebogen und in Diskussionen erhoben und in qualitativer Inhaltsanalyse nach Mayring [37] analysiert wurden. Jeder Vorsatz wird einer Kategorie (Individuum, Organisation, IT-Infrastruktur) zugeordnet, abhängig worauf sich ein Vorsatz auswirkt (siehe Tabelle 5).

Tabelle 5. Ausgewählte Vorsätze der Spielteilnehmer im Hinblick auf eine Verhaltensänderung

Individuum	*Organisation*	*IT-Infrastruktur*
Büros nicht unverschlossen lassen; Mehr Berichte / TecBlogs zum Thema IT-Sicherheit / Cyberwar lesen; Mehr Zeit in IT-Sicherheit investieren; Bewusster auf potentielle Gefahren achten;	Transparenz in der IT-Sicherheit; Sensibilisierung der Mitarbeiter intensivieren; Mit den anderen Administratoren Angriffsszenarien ausdenken und dann checken, ob man safe ist; Kompetenzen auf mehrere Schultern verteilen;	Öfter eigenes Netzwerk penetrieren; Regelmäßige Schwachstellenanalyse; Datenabfluss über Web-Schnittstellen (Cloud, Email, …) verhindern / eindämmen; Testumgebung für kommende Updates nutzen (Office, Windows 7, Java etc.);

Auffällig ist, dass viele Vorsätze genannt werden, die die Organisation betreffen: Organisation (117 Vorsätze), Individuum (78 Vorsätze), IT-Infrastruktur (13 Vorsätze). Das zeigt, dass „Operation Digitales Chamäleon“ eher organisatorische IT-Sicherheit als technischen Schutz der IT-Infrastruktur adressiert. Gerade vor dem Hintergrund, dass sich im Themenfeld IT-Sicherheit Kritischer Infrastrukturen IT-Professionals um die Sicherheit von bspw. Industrieanlagen Gedanken machen müssen, die bisher nicht als IT-Sicherheitsrisiko eingeschätzt wurden, kann ein solcher Impuls die notwendigen organisatorischen Änderungen anregen.

Tabelle 6. Kategorien und Anzahl der Vorsätze für Verhaltensänderung

Kategorie	*Anzahl der Vorsätze*
(Organisatorisches) IT-Security Awarenesstraining	55
(Individuelles) IT-Security Awarenesstraining	27
Operative IT-Security-Vorgänge	19
Monitoring	18
Bestehende IT-Security-Konzepte	15
Aufgeschlossene Denkansätze	14
Cross-Functional IT-Security-Teams	13
Aufmerksamkeit	11
Informationsverteilung	10
Informationsbeschaffung	6
Auditing	5
Penetration Testing	5
Abschreckung	3
Personalmanagement	3
Resignation	2
Keine Verhaltensänderung	1
Keine Auswertung möglich	1

In einer qualitativen Inhaltsanalyse wurden die 208 Vorsätze in 17 Kategorien systematisiert (vgl. Tabelle 6). 55 Vorsätze werden der Kategorie „(Organisatorisches) IT-Security Awarenesstraining“ zugeordnet: Spielteilnehmer möchten die Mitarbeiter der Organisation besser sensibilisieren. 27 Vorsätze umfasst die Kategorie „(Individuelles) IT-Security Awarenesstraining“, mit Vorsätzen wie „Erhöhte Weiterbildung im Bereich IT-Sicherheit“. Kategorie „Operative IT-Security-Vorgänge“ beinhaltet 19 Vorsätze – ein Beispiel ist, zu prüfen, „Dass keine unbekannte Person sich alleine im Raum aufhält“. Auch dies illustriert, dass das Spiel Impulse setzt, IT-Sicherheit als Organisationsaufgabe zu begreifen.

In den Spieldurchführungen konnte beobachtet werden, dass interdisziplinäre Teams gute Resultate erzielen, also Teams mit Mitgliedern entweder verschiedener Hierarchiestufen von der strategischen bis hin zur operativen Ebene oder mit Mitgliedern aus den Kernthemen der IT-Sicherheit zusammen mit Vertretern angrenzender Domänen wie Arbeitsschutz oder bauliche Sicherheit. So nutzte bspw. ein blaues (interdisziplinäres) Team Motivatoren zur Steigerung der Mitarbeiterzufriedenheit um die Bereitschaft der Mitarbeiter, Innentäter zu werden zu reduzieren. Die Notwendigkeit für Interdisziplinarität spiegelt sich auch in den Vorsätzen wieder. So beinhaltet Kategorie „Cross-Functional IT-Security-Teams“ 13 Vorsätze, von denen ein Vorsatz lautet: „Zusammenarbeit mit anderen (Admins, User, SiBe, Brandschutzbeauftragter)“.

Nur vergleichsweise wenige Spielteilnehmer planen als Vorsatz die IT-Sicherheitstechnologie zu verbessern oder zu erneuern. Spielteilnehmer gaben an, sie wollen bspw. häufiger nach Schwachstellen oder Fehlkonfigurationen in der IT-Infrastruktur suchen und diese beheben („Penetration Testing“ (5 Vorsätze)).

Zusammenfassend lässt sich feststellen, dass „Operation Digitales Chamäleon“ zu Verhaltensänderungen motiviert, also nicht alleine Faktenwissen vermittelt und vor allem zur Auseinandersetzung mit organisatorischen Aspekten anregt.

6 Zusammenfassung und Ausblick

In diesem Artikel präsentierten wir unsere IT-Security Awarenessmaßnahme „Operation Digitales Chamäleon“ mit ausgewählten Ergebnissen der ersten sieben Durchführungen. Die APTs als Spielergebnisse zeigen, dass die Spielteilnehmer APTs und IT-Sicherheitsmaßnahmen realistischer Komplexität erarbeiten und so „Operation Digitales Chamäleon“ geeignet ist, als IT-Security Awarenessmaßnahme für IT-Sicherheitsverantwortliche und IT-Sicherheitsprofessionals im KRITIS Kontext eingesetzt zu werden (vgl. [5] [6]). Die Auswahl der Ergebnisse liegt in diesem Beitrag auf dem Spielerlebnis, der Wahrnehmung, dem Wissensgewinn und der Verhaltensänderung. Das Spiel macht Spaß und erfüllt damit ein wichtiges Erfolgsmerkmal von Serious Games. „Operation Digitales Chamäleon“ motiviert die Spielteilnehmer zu Verhaltensänderungen, und vermittelt nicht alleine Faktenwissen. Diese Verhaltensänderungen betreffen eher organisatorische als individuelle oder technische Aspekte. „Operation Digitales Chamäleon“ adressiert also die wichtigen Themen der IT-Sicherheit Kritischer Infrastrukturen. Das deckt sich mit den Empfehlungen des BSI: Langfristige Verhaltensänderungen sind wesentlich für ein langfristig erfolgreiches IT-Sicherheitsmanagement – neue IT-Sicherheitstechnologie anzuschaffen genügt nicht. Die Evaluation des Erfolgs solcher IT-Sicherheitsawarenessmaßnahmen ist inhärent schwierig zu evaluieren – das zeigen der State of the Art von Praxis und wissenschaftlicher Literatur und so ist die Evaluation des Spieles dem Spielgegenstand angemessen.

„Operation Digitale Schlange“ und „Operation Digitale Eule“ sind zusammen mit „Operation Digitales Chamäleon“ die ersten Spielformate der IT-Security Matchplay Serie. Für alle drei Formate sind weitere Spiele und Weiterentwicklungen geplant.

7 Danksagung

Hiermit bedanken wir uns beim BMBF als Fördergeber des Projekts „Vernetzte IT-Sicherheit Kritischer Infrastrukturen“ (FKZ 16KIS0213). Ebenso möchten wir allen Spielteilnehmern für die guten Spielergebnisse und das Engagement danken. Dem Associate Editor und den Gutachtern sind wir für hilfreiche und konstruktive Kommentare dankbar.

Literaturverzeichnis

1. McAfee: Combating Advanced Persistent Threats. , Santa Clara (2011).
2. Symantec: Advanced Persistent Threats: How They Work, http://www.symantec.com/theme.jsp?themeid=apt-infographic-1.
3. Rowney, K.: What We Talk About When We Talk About APT, http://www.symantec.com/connect/blogs/what-we-talk-about-when-we-talk-about-apt#!
4. Rouse, M.: advanced persistent threat (APT), http://searchsecurity.techtarget.com/definition/advanced-persistent-threat-APT.
5. Rieb, A., Lechner, U.: Operation Digital Chameleon – Towards an Open Cybersecurity Method. In: Proceedings of the 12th International Symposium on Open Collaboration

(OpenSym 2016). pp. 1–10. , Berlin (2016).
6. Rieb, A., Lechner, U.: Towards Operation Digital Chameleon. In: Havârneanu, G., Setola, R., Nassopoulos, H., and Wolthusen, S. (eds.) CRITIS 2016 - The 11th International Conference on Critical Information Infrastructures Security (to appear). pp. 1–6. , Paris (2016).
7. Wilson, M., Hash, J.: NIST Special Publication 800-50: Building an Information Technology Security Awareness and Training Program. , Gaithersburg (2003).
8. ENISA: Der neue Leitfaden für die Praxis: Wege zu mehr Bewusstsein für Informationssicherheit. Europäische Agentur für Netz- und Informationssicherheit (ENISA) (2008).
9. ENISA: Information security awareness initiatives: Current practice and the measurement of success. (2007).
10. BSI: Awareness-Umfrage 2015, https://www.allianz-fuer-cybersicherheit.de/ACS/DE/_/downloads/awareness-umfrage-2015.pdf?__blob=publicationFile&v=5, (2016).
11. BSI: IT-Grundschutz: Glossar und Begriffsdefinitionen, https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKataloge/Inhalt/Glossar/glossar_node.html.
12. BSI: ERFA-Kreis Awareness, https://www.allianz-fuer-cybersicherheit.de/ACS/DE/Erfahrungsaustausch/ERFA-Kreise/Awareness/awareness.html.
13. BSI: IT-Grundschutz - B 1.13 Sensibilisierung und Schulung zur Informationssicherheit, https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKataloge/Inhalt/_content/baust/b01/b01013.html.
14. BSI: IT-Grundschutz - M 3.47 Durchführung von Planspielen zur Informationssicherheit, https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKataloge/Inhalt/_content/m/m03/m03047.html.
15. Blötz, U.: Planspiele und Serious Games in der beruflichen Bildung: Auswahl, Konzepte, Lernarrangements, Erfahrungen - Aktueller Katalog für Planspiele und Serious Games (Berichte zur beruflichen Bildung). W. Bertelsmann Verlag GmbH & Co. KG, Bielefeld (2015).
16. Schwägele, S.: Planspiel - Lernen - Lerntransfer. Eine subjektorientierte Analyse von Einflussfaktoren, file:///home/arieb/Downloads/SchwaegeleDissopusse_A3a.pdf.
17. Hansch, N., Benenson, Z.: Specifying IT security awareness. In: Proceedings - International Workshop on Database and Expert Systems Applications, DEXA. pp. 326–330. , München (2014).
18. Bulgurcu, B., Cavusoglu, H., Benbasat, I.: Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness. MIS Q. 34, 523–548 (2011).
19. Johnston, A.C., Warkentin, M.: Fear Appeals and information Security Behaviors: An Empirical Study. MISQ. 34, 549–566 (2010).
20. Furnell, S., Moore, L.: Security literacy: The missing link in today's online society? Comput. Fraud Secur. 2014, 12–18 (2014).
21. Perla, P.P.: The Art of Wargaming: A Guide for Professionals and Hobbyists. US Naval Institute Press (1990).
22. ENISA: On National and International Cyber Security Exercises. Europäische Agentur für Netz- und Informationssicherheit (ENISA), Heraklion (2012).
23. ENISA: The 2015 Report on National and International Cyber Security Exercises. Europäische Agentur für Netz- und Informationssicherheit (ENISA), Athen (2015).
24. Cetto, A., Netter, M., Pernul, G.: Friend Inspector: A Serious Game to Enhance Privacy

Awareness in Social Networks. Proc. 2nd Int. Work. Intell. Digit. Games Empower. Incl. (IDGEI '13). 1–8 (2014).

25. Netter, M., Pernul, G., Richthammer, C., Riesner, M.: Privacy in Social Networks: Existing Challenges and Proposals for Solutions. Commun. Comput. Inf. Sci. 576, 16–27 (2015).
26. Shostack, A.: Elevation of Privilege: Drawing Developers into Threat Modeling. USENIX Summit Gaming, Games, Gamification Secur. Educ. 1–15 (2014).
27. Beckers, K., Pape, S.: A Serious Game for Eliciting Social Engineering Security Requirements. Presented at the (2016).
28. PWC: Game of ThreatsTM – Cybersecurity-Simulation für Manager, http://www.pwc.de/de/digitale-transformation/game-of-threats-cybersecurity-simulation-fuer-manager.html.
29. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Q. 28, 75–105 (2004).
30. Baskerville, R., Pries-Heje, J.: Explanatory Design Theory. Bus. Inf. Syst. Eng. 2, 271–282 (2010).
31. Hald, S., Pedersen, J.: An updated taxonomy for characterizing hackers according to their threat properties. Adv. Commun. Technol. (ICACT), 2012 14th Int. Conf. 81–86 (2012).
32. BSI: Industrie und Kritische Infrastrukturen: Meldepflicht, https://www.bsi.bund.de/DE/Themen/Industrie_KRITIS/IT-SiG/Neuregelungen_KRITIS/Meldepflicht/meldepflicht_node.html.
33. Kriz, W.C., Nöbauer, B.: Den Lernerfolg mit Debriefing von Planspielen sichern, http://www.bibb.de/dokumente/pdf/1_08a.pdf.
34. Kriz, W.C., Hense, J.: Qualitätskriterien von Planspielprodukten. In: Blötz, U. (ed.) Planspiele und Serious Games in der beruflichen Bildung2. pp. 222–223. W. Bertelsmann Verlag GmbH & Co. KG, Bielefeld (2015).
35. McConigal, J.: Besser als die Wirklichkeit!: Warum wir von Computerspielen profitieren und wie sie die Welt verändern. Heyne Verlag, München (2012).
36. Killmeyer, J.: Information Security Architecture: An Integrated Approach to Security in the Organization. Auerbach Publications, Boca Raton, New York (2006).
37. Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. (2008).
38. BSI: IT-Grundschutz - M 2.73 Auswahl geeigneter Grundstrukturen für Sicherheitsgateways, https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKataloge/Inhalt/_content/m/m02/m02073.html.
39. heise: Datendiode gegen Datendiebe, http://www.heise.de/newsticker/meldung/Datendiode-gegen-Datendiebe-2139499.html.
40. Shulman, A., Dulce, S.: Man in the Cloud (MITC) Attacks. (2015).
41. Gragido, W.: Lions At The Watering Hole - The "VOHO" Affair, http://blogs.rsa.com/lions-at-the-watering-hole-the-voho-affair/.

Entwurf eines cloud-basierten Geschäftsmodells für die kontinuierliche Prüfung

Andreas Kiesow[1], Oliver Thomas[1]

[1] Universität Osnabrück, Fachgebiet für Informationsmanagement und Wirtschaftsinformatik, Osnabrück, Deutschland
{andreas.kiesow,oliver.thomas}@uni-osnabrueck.de

Abstract. Die zunehmende Digitalisierung und Komplexität der Abschlussprüfung als daten- und wissensintensive Dienstleistung externer Prüfungsgesellschaften verändert bestehende Geschäftsabläufe der Mandanten-Prüfer-Beziehung. Dabei ist ein Wandel von einer jährlichen Prüfung auf Basis historischer Daten zu einer unterjährigen, kontinuierlichen Prüfung zu beobachten. Daraus ergibt sich, insbesondere für kleine und mittelständische Prüfungsgesellschaften, der Bedarf an technologischer Unterstützung, um ihre Stellung am Abschlussprüfermarkt zu behaupten. Ein innovativer Lösungsansatz ist in diesem Zusammenhang die Auslagerung kontinuierlicher Prüfungsdienstleistungen an einen spezialisierten Informationsdienstleister, der Prüfungsresultate über eine Cloud-Architektur bedarfsorientiert zur Verfügung stellt (Audit-as-a-Service). Zur Umsetzung dieses Ansatzes wird in dem vorliegenden Artikel ein Geschäftsmodell nach dem Business Model Canvas entworfen. Dazu wurde ein konzeptionell-deduktives Verfahren angewendet. Die Autoren dieses Artikels adressieren mit der Vorstellung dieses Geschäftsmodellentwurfs eine Forschungslücke in der Fachliteratur und eröffnen damit eine neue Forschungsrichtung zur Umsetzung und Verbreitung von kontinuierlichen Prüfungsdienstleistungen.

Keywords: Continuous Auditing, Audit-as-a-Service, Business Model, Cloud-Computing, Digitalisierung

1 Einleitung

Geschäftsmodelle (engl. business models) können als Werkzeuge verstanden werden, die die Beschreibung, Erneuerung und Bewertung von Geschäftsabläufen ermöglichen. Geschäftsmodelle werden als *digital* bezeichnet, wenn eine Veränderung der genutzten Informations- und Kommunikationstechnik (IKT) fundamentale Veränderungen für die Durchführung des Geschäfts und die Umsatzentstehung nach sich ziehen [1]. Eine zunehmende Digitalisierung der Geschäftsabläufe ist bei der Prüfung von Rechnungslegungsprozessen im Rahmen der Jahresabschlussprüfung zu beobachten. Die Verschärfung regulatorischer Rahmenbedingungen und steigende technische Komplexität der Rechnungslegung erfordern einen erheblichen Bedarf zur Weiterentwicklung traditioneller Geschäftsabläufe in der Abschlussprüfung [2], [3].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Kiesow, A.; Thomas, O. (2017): Entwurf eines cloud-basierten Geschäftsmodells für die kontinuierliche Prüfung, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 882-896

Dieser Wandel ist vorrangig durch eine Veränderung von einer traditionellen Jahresabschlussprüfung mit Prüfungshandlungen auf Basis historischer Daten in Stichproben [4] zu einer unterjährigen, automatisierten Prüfung in (quasi-)Echtzeit, gekennzeichnet [5]. Diese kontinuierliche Prüfung (engl. continuous auditing) basiert auf dem Einsatz innovativer Technologien und der Entwicklung betrieblicher Informationssysteme (IS), deren Untersuchung einen Kernbereich der Wirtschaftsinformatik (WI) darstellt [6]. Trotz der Implementierung von Prototypen in individualisierten Einzelszenarien, wie z. B. [7], [8], haben sich kontinuierliche Prüfungsansätze in der Praxis bislang nicht verbreitet [9]. Als Gründe werden dafür technische und organisatorische Barrieren [10] sowie mangelnde fachliche Expertise [11] genannt. In der Fachliteratur existieren in diesem Kontext Überlegungen hinsichtlich der Auslagerung von kontinuierlichen Prüfungsdienstleistungen an spezialisierte Informationsdienstleister, um insbesondere kleinen und mittelständischen Wirtschaftsprüfungsgesellschaften (WPG) Prüfungsergebnisse auf kontinuierlicher Basis zugänglich zu machen [12], [13].

Durch die mittlerweile nahezu durchgängige Digitalisierung der Rechnungslegung durch Enterprise-Resource-Planning-, und Buchhaltungssysteme [14] sowie die jederzeitige Verfügbarkeit und Flexibilität von Ressourcen durch Virtualisierung sind wesentliche technische Voraussetzungen zur Auslagerung kontinuierlicher Prüfungsdienstleistungen erfüllt. In dem vorliegenden Artikel wird eine solche Auslagerung von Prüfungshandlungen und den damit verbundenen vor- und nachgelagerten Tätigkeiten an einen externen Informationsdienstleister, der als Bindeglied zwischen Mandanten und Prüfungsgesellschaft fungiert, betrachtet.

Zur Umsetzung dieses Auslagerungsvorhabens bedarf es allerdings Geschäftsmodelle als Werkzeug zur Repräsentation und Analyse der angestrebten Geschäftslogik [15], [1]. Der vorliegende Beitrag greift diesen Bedarf an Geschäftsmodellen für die kontinuierliche Prüfung auf. Dazu werden die Eigenschaften des Cloud-Computings (CC) zu Entwicklung einer Geschäftsmodellinnovation *Audit-as-a-Service* genutzt. Da diese Umsetzung die Entwicklung eines digitalen Geschäftsmodells erfordert [16], lautet die Forschungsfrage dieses Artikels daher: *Wie muss das cloud-basierte Geschäftsmodell eines externen Dienstleisters für die kontinuierliche Prüfung von Rechnungslegungsprozessen gestaltet werden, um die Verbreitung der kontinuierlichen Prüfung in der Praxis zu ermöglichen?*

Zur Beantwortung dieser Frage wird auf Grundlage des Business Model Canvas nach Osterwalder (2004) und Osterwalder und Pigneur (2010) ein Geschäftsmodell entworfen und unter Anwendung des Ordnungsrahmens nach Labes et al. (2014) hinsichtlich cloud-spezifischer Bewertungskriterien analysiert. Der vorliegende Artikel stellt erstmalig in der Fachliteratur den Entwurf eines vollständigen Geschäftsmodells für die kontinuierliche Prüfung vor. Damit werden die bislang vorrangig technologischen Betrachtungen der digitalen Transformation der Abschlussprüfung um Aspekte der Geschäftsmodellentwicklung erweitert. Das entworfene Geschäftsmodell kann sowohl von Wissenschaftlern als auch Praktikern als Werkzeug zur Analyse der Voraussetzungen und Aktivitäten, die für die Umsetzung und Verbreitung der kontinuierlicher Prüfung erforderlich sind, genutzt werden.

Dazu wurde der Artikel wie folgt gegliedert: In Abschnitt 2 werden relevante theoretische Grundlagen zur kontinuierlichen Prüfung und zur Geschäftsmodellentwicklung erläutert. In Abschnitt 3 wird das durchgeführte Forschungsvorgehen begründet. In Abschnitt 4 werden die Ergebnisse zur Geschäftsmodellentwicklung und -analyse gezeigt. Die Ergebnisse und Limitationen des Artikels werden in Abschnitt 5 kritisch reflektiert.

2 Theoretische Grundlagen

2.1 Kontinuierliche Prüfung

Um eine fundierte Grundlage über den Forschungsstand der kontinuierlichen Prüfung zu bilden, wurde eine strukturierte Literaturanalyse nach Webster und Watson (2002) durchgeführt [17]. Bei der Suche in sechs Literaturdatenbanken (EbscoHost, SpringerLink, ISI Web of Knowledge, ACM Digital Library, ScienceDirect und Wiley) nach den Suchbegriffen {Continuous Assurance}, {Continuous Auditing} und {Continuous Monitoring + Audit} wurden 52 relevante Artikel identifiziert, die durch Rückwärts- und Vorwärtssuche um weitere 42 Artikel ergänzt wurden. Die Relevanzkriterien waren dabei englische Sprache, Qualität des Publikationsorgans, die Anwendung im Kontext der Jahresabschlussprüfung (engl. financial auditing) und Verfügbarkeit des Artikels. Die Kriterien zur Analyse und Charakterisierung der 94 Papiere waren der grundlegende wissenschaftliche Ansatz, die betrachteten technischen Konzepte zur Umsetzung der kontinuierlichen Prüfung, die Zuordnung zur internen und/oder externen Revision sowie das Artefakt des Papiers, wie z. B. Methoden oder Modelle. Die detaillierte Beschreibung und zentralen Ergebnisse dieser Literaturanalyse wurden als Vorarbeiten in einem separaten Papier [18] veröffentlicht. Für den vorliegenden Artikel wurde die Analyse um die Suche nach dem Begriff „business model“ erweitert und die so identifizierten acht Artikel ausgewertet.

Nachweise zu Überlegungen, digitale Geschäftsmodelle der Mandanten für eine kontinuierliche Prüfung der Rechnungslegungsdaten zu nutzen, lassen sich in der Fachliteratur seit Verbreitung des Internets zu Beginn der 2000er Jahre. Kneer (2003) stellt dazu die Frage: „Has the change in the business model/environment rendered look-back financial attestations valueless?“ [19]. Greenstein und Ray (2002) stellen in diesem Kontext ein Modell zur Integration von Prüfungsleistungen in die Geschäftsprozesse der Mandanten vor: „This model is based on accounting firms' need to respond quickly to evolving e-business models employed by their information-age clients“ [20]. Murthy und Groomer (2004) sehen in der kontinuierlichen Prüfung die Möglichkeit zur Entwicklung eines innovativen Geschäftsmodells und stellen ein Modell vor, das den bedarfsorientierten Zugriff auf Prüfungsergebnisse und -urteile gegen eine Gebühr, vorsieht [21]. Yeh und Shen (2010) erweitern dieses Modell um technologische Aspekte und kommen zu dem Ergebnis, dass darauf ein neues Geschäftsmodell für Wirtschaftsprüfer und die externe Revision aufgebaut werden könnte [22]. Gehrke und Wolf (2010) stellen das Konzept einer Web 2.0 Kooperationsplattform für Abschlussprüfer vor und skizzieren in diesem Zusammenhang ein Geschäftsmodell für die kontinuierliche Prüfung.

Bei allen vorgenannten Autoren wird stets von einer Prüfer-Mandanten-Beziehung ausgegangen. Ein alternativer Ansatz ist die Auslagerung von Prüfungshandlungen und den damit verbundenen vor- und nachgelagerten Tätigkeiten an einen externen Informationsdienstleister, der als Bindeglied zwischen Mandanten und WPG fungiert [12]. Der Vorteil für die WPG liegen darin, dass sie nicht selbst die personellen Kapazitäten, die für die Umsetzung der kontinuierlichen Prüfung erforderlich sind, aufbauen müssen, da der Informationsdienstleister diese als temporäre Leistung der WPG zur Verfügung stellt. Die WPG profitiert zudem von den Erfahrungen, die ein Informationsdienstleister aufbauen wird, dessen Tagesgeschäft die kontinuierliche Prüfung ist. Begünstigt durch Erfahrungskurven- und Skaleneffekte ist davon auszugehen, dass die Kosten der Auslagerung niedriger sind als im eigenen Kompetenzaufbau. Der Informationsdienstleister wird langfristig Standards bei den Anbindungen, Extraktion und Auswertung der Rechnungslegungsdaten entwickeln, so dass eine zeitnahe Anpassung der Extraktions- und Auswertungsmechanismen an heterogene Systemlandschaften angenommen werden kann.

Allerdings wurden im Rahmen der durchgeführten Literaturanalyse keine detaillierten Ausführungen zu Geschäftsmodellinhalten eines solchen Ansatzes gefunden. Zusammengefasst werden zwar die Notwendigkeit und die technische Machbarkeit von Geschäftsmodellen zur Verbreitung kontinuierlicher Prüfung gesehen [15], jedoch keine fundierten Untersuchungen über die Gestaltung, Elemente und die Wertschöpfung der kontinuierlichen Prüfung durchgeführt. Diese Forschungslücke wird mit dem vorliegenden Artikel adressiert.

2.2 Geschäftsmodelle

In der Fachliteratur existieren unterschiedliche Definitionen des Begriffs Geschäftsmodell [23]. Morris et al. (2005) definieren das Geschäftsmodell als prägnante Darstellung von untereinander verbundenen Variablen, deren Verbesserung einen nachhaltigen Wettbewerbsvorteil implizieren [24]. Osterwalder (2004) sehen im Geschäftsmodell eine vereinfachte Darstellung der Aktivitäten eines Unternehmens. Dabei wird das Geschäftsmodell als Werkzeug oder Instrument charakterisiert. Der Zweck eines Geschäftsmodells besteht demnach darin, die Kombination der benötigten Elemente und deren Beziehungen untereinander aufzuzeigen [25]. Nach Amit und Zott (2001) sind diese Elemente die Gestaltung der Inhalte, des Aufbaus und der Steuerung von Geschäftsvorgängen [26]. Nach Bieger und Reinhold (2011) sowie Chesbrough und Rosenbloom (2002) kann der Inhalt eines Geschäftsmodells auch als Logik des Unternehmens verstanden werden [27], [28]. Alle Definitionen teilen das Verständnis, dass die zentrale Funktion des Geschäftsmodells die Darstellung des Nutzens oder Wertes, der von der Organisation vermittelt wird, beschreibt [25–28]. Dieser Nutzen fließt dabei an Kunden, Partner und an das Unternehmen selbst [26], [25]. Auf Basis dieser Beschreibungen wird in diesem Artikel das Geschäftsmodell definiert, wie folgt: Ein Geschäftsmodell ist ein Instrument zur präzisen Darstellung der Logik, bzw. der interagierenden Elemente einer Unternehmung, um Nutzen für Kunden, Partner und sich selbst zu schaffen und dabei einen dauerhaften Wettbewerbsvorteil zu erlangen.

Das Geschäftsmodell dient in diesem Zusammenhang als Medium zwischen Geschäftsstrategie und den Geschäftsprozessen des Unternehmens [1].

In der Fachliteratur werden unterschiedliche Konzepte zur Erstellung von Geschäftsmodellen diskutiert, die sich insbesondere hinsichtlich der zu berücksichtigenden Elemente unterscheiden [23], [29]. Auf Basis vergleichender Literatur wurden für die Untersuchungen zu diesem Artikel ausschließlich Konzepte betrachtet, die ihre jeweiligen Elemente explizit benennen [23, 30]. In den Vorarbeiten für den vorliegenden Artikel, die sich hier nicht darstellen lassen, wurden insgesamt sieben Konzepten hinsichtlich relevanter Elemente analysiert. Die Geschäftsmodell-konzepte nach Osterwalder (2004), Bieger und Reinhold (2011) sowie Hamel (2002) wurden dabei als die geeignetsten Konzepte hinsichtlich Abdeckungsrad relevanter Elemente bewertet. Diese drei Konzepte sind zudem allgemeingültig und unabhängig einer bestimmten Industrie anwendbar [31], [30].

Die Autoren dieses Papiers entschieden sich für die Anwendung des Business Models Canvas nach Osterwalder (2004) sowie Osterwalder und Pigneur (2010). Dieser Ansatz ist dem strategischen Management und dem Informationsmanagement zuzuordnen, sowohl in der Praxis als auch in der Wissenschaft etabliert und durch eine intuitive Vorgehensweise gekennzeichnet [30]. Anders als bei Hamel (2002) und Bieger und Reinhold (2011) beginnt bei Business Model Canvas der Prozess eines innovativen Geschäftsmodells bei der Recherche, Gestaltung und Implementierung eines neuen Geschäftsmodells, welches anhand der Marktreaktionen weiterentwickelt wird. Aus Sicht der Autoren ist dieser Ansatz für eine gestaltungsorientierte Entwicklung eines Geschäftsmodells gegenüber den anderen Ansätzen zu bevorzugen.

3 Forschungsansatz

Nach Magretta (2002) kann das Geschäftsmodell als praktisches Äquivalent einer Forschungsmethode verstanden werden: “Business modeling is the managerial equivalent of the scientific method–you start with a hypothesis, which you then test in action and revise when necessary” [32]. Dies impliziert das Vorhandensein eines Geschäftsmodells, das in anknüpfenden Forschungsarbeiten zum Test, d. h. zur empirischen Evaluation, verwendet und zukünftig verbessert werden kann. Die Gestaltung eines Geschäftsmodells für die kontinuierliche Prüfung erfordert aus Sicht der Autoren daher ein konstruktivistisches Vorgehen. Diese Notwendigkeit begründet sich aus dem grundsätzlichen Verständnis des Geschäftsmodells als Artefakt, mit dem die kontinuierliche Prüfung effektiv und effizient umgesetzt werden kann [33].

Bei der Abwägung von Methoden sind aus Sicht der Autoren wirtschaftliche und zeitliche Einschränkungen zu berücksichtigen. Simulationen, prototypische Umsetzungen und Methoden der Aktionsforschung erfordern kapitalintensive Vorarbeiten, die aufgrund der Komplexität der Abschlussprüfung und den dabei involvierten Akteuren nur unter hohem Kapitaleinsatz zu entwickeln sind. Zum gegenwärtigen Zeitpunkt ist aus Sicht der Autoren der initiale Entwurf eines Geschäftsmodells für die Auslagerung von Prüfungsleistungen einzig durch die deduktive Ableitung eines Konzepts mit Referenzeigenschaft auf Planniveau möglich

[34]. Dieser initiale Entwurf kann für die Entwicklung zukünftiger IKT in der Wirtschaftsprüfung verwendet werden, um, wie von Magretta gefordert, einen empirischen Nachweis über die Nützlichkeit der kontinuierlichen Prüfung in der Praxis zu erbringen.

Gestaltungsorientierte Forschung (engl. design science research) erfordert die rigorose Evaluation der entwickelten Artefakte [33]. Aufgrund der komplexen, praktischen Umsetzung des in diesem Artikel diskutierten, neuartigen Prüfungsansatzes als kapitalintensive Innovationen ist die Durchführung von Evaluationsmethoden höherer Wertigkeit, wie z. B. Fallstudien oder Simulationen, zum gegenwärtigen Zeitpunkt nicht möglich. Entsprechend der konzeptionellen Ausrichtung des vorliegenden Artikels, kann die Evaluation des vorgestellten Geschäftsmodells derzeit ausschließlich rein deskriptiv erfolgen. Dazu wurde eine argumentativ-deduktive Analyse unter Verwendung des morphologischen Kastens nach Labes et al. 2014 durchgeführt [16]. Zusammengefasst wurde in der vorliegenden Forschungsarbeit ein konstruktivistisch-qualitatives Vorgehen angewendet.

4 Geschäftsmodell Audit-as-a-Service

4.1 Kurzbeschreibung Audit-as-a-Service

Audit-as-a-Service (AaaS) umfasst die Integration eines spezialisierten Informationsdienstleisters als Bindeglied zwischen Wirtschaftsprüfungsgesellschaften (WPG) und deren Mandanten und stellt somit eine vertikale Erweiterung der Wertschöpfungskette in der Wirtschaftsprüfung dar (siehe Abbildung 1). Die Dienstleistung des Informationsdienstleisters ist ein Leistungsbündel aus einmaliger Implementierung von Extraktionsroutinen in den Rechnungslegungssystemen des Mandanten, z. B. Buchhaltungssystemen, und der kontinuierlichen Übertragung und Auswertung der daraus extrahierten Daten. Kerngedanke ist somit, dass die WPG die kontinuierliche Prüfung und die dazu erforderlichen Aktivitäten an den Informationsdienstleister auslagert [35]. Der WPG soll die Möglichkeit gegeben werden, die Ergebnisse dieser Prüfung bedarfsgerecht, d. h. in eigens vordefinierten Intervallen und nach individualisierten Auswertungsregeln, zu beziehen. An diese besondere Form der Auslagerung werden verschiedene Anforderungen hinsichtlich Bedarfs- und Serviceorientierung gestellt. Die Eigenschaften des Cloud-Computings (CC), maßgeblich das bedarfsgerechte und flexible Angebot von IT-Leistungen in Echtzeit, bieten anerkanntermaßen ein hohes Potenzial zur Entwicklung disruptiver Geschäftsmodellinnovationen [36].

Nachfolgend wird der initiale Entwurf eines Geschäftsmodells vorgestellt, das die Eigenschaften des Cloud-Computings mit dem Prinzip der kontinuierlichen Prüfung zu einem Geschäftsmodell „as-a-Service" vereint. Diese Ausführungen basieren auf Vorarbeiten, die die Untersuchung von Erfolgsfaktoren für die kontinuierliche Prüfung auf Basis einer Literaturstudie [10], die Entwicklung eines Rahmenwerks zur Gestaltung von IS für die kontinuierliche Prüfung auf Basis von Fokusgruppeninterviews mit Experten aus der IT-gestützten Abschlussprüfung und dem Enterprise Architecture Management [35] sowie ein im Anschluss erstelltes Prozessmodell für Audit-as-a-Service, das auf Basis von Eins-zu-Eins-Diskussionen mit Experten aus der Praxis erstellt wurde [37].

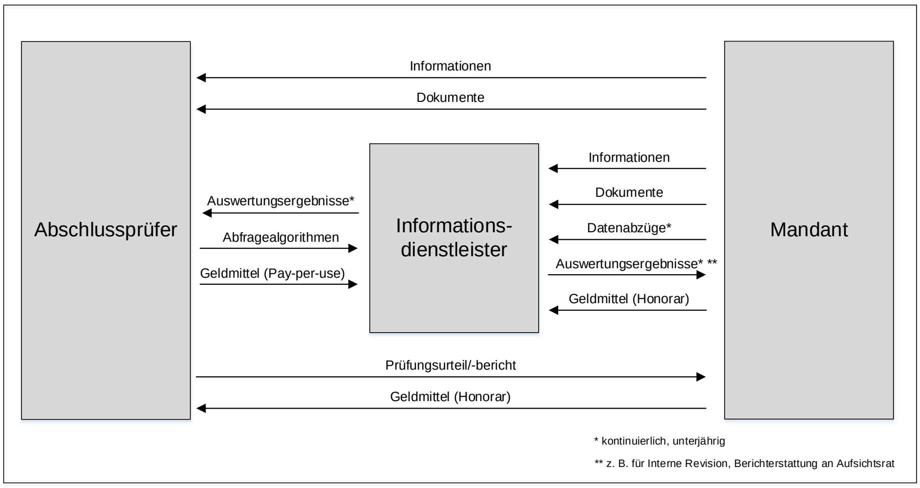


Abbildung 1. Wertschöpfungskette von Audit-as-a-Service [13]

4.2 Geschäftsmodellkonzeption

Das Business Model Canvas nach Osterwalder (2004) besteht aus neun unterschiedlichen Elementen [25]. Nachfolgend werden diese neun Elemente für AaaS beschrieben. Eine zusammenfassende Darstellung dieser neun Elemente ist entsprechend des Business Models Canvas [38] in Abbildung 2 enthalten.

(1) Kundensegmente. Die Kundensegmente sind der erste Teil der Kundenschnittstelle und legen fest, an welche Zielkundengruppe sich das Wertversprechen richtet [25]. Kunden von AaaS sind sowohl Organisationen, die eine Rechnungslegung betreiben und ein Interesse an einem kontinuierlichen Nachweis über die Vollständigkeit und Richtigkeit der verarbeiteten Rechnungslegungsdaten haben als auch Organisationen, die im Auftrag die Rechnungslegung einer anderen Organisation prüfen. Dabei werden vorrangig kleine und mittelständische WPG adressiert, deren personellen und technischen Kapazitäten häufig nicht ausreichen, um selbst eine kontinuierliche Prüfung umzusetzen. Zum erweiterten Kundenkreis zählen Abteilungen der Internen Revision, Controlling und Business Intelligence von Großunternehmen sowie Start-Ups und Organisationen im öffentlichen Sektor.

(2) Wertversprechen. Das Wertversprechen ist das zentrale Element des Geschäftsmodells und repräsentiert das Wertangebot des Unternehmens. Dabei handelt es sich um ein Bündel aus Produkten oder Dienstleistungen, die einen Nutzen für die Kunden schaffen und auf den Ressourcen des Unternehmens basieren [25]. Nach

Osterwalder und Pigneur (2010) schaffen verschiedene qualitative und quantitative Faktoren einen Wert für Kunden [38]. Zielsetzung von AaaS ist die Befähigung der Kunden zur kontinuierlichen Prüfung von Rechnungslegungsprozessen. Das Attribut kontinuierlich bedeutet dabei *unterjährig* und *zeitnah*, was die Automatisierung und Digitalisierung von Prüfungshandlungen erfordert. Die *Prüfung* beschreibt den Vorgang der Beurteilung, ob die Daten der Rechnungslegung einer Organisation ein den tatsächlichen Verhältnissen entsprechendes Bild der Vermögens-, Finanz- und Ertragslage wiedergeben. Damit soll der Zeitverzug zwischen Entstehung und Aufdeckung eines inkorrekten oder betrügerischen Sachverhaltes reduziert werden. Zudem sollen die üblichen Belastungsspitzen in den Jahresabschlüssen reduziert und die Qualität der Rechnungslegungsdaten der Organisation erhöht werden. Betriebszweck von AaaS ist es, den Kunden diesen Nutzen zu bieten, ohne dass diese eigenen, kapitalintensiven Anschaffungen, wie z. B. Hard- und Software sowie personelle Kapazitäten, aufbauen müssen. Dieses Wertversprechen wird durch den Aufbau einer Programmier- und Software-Umgebung (Platform-as-a-Service, Software-as-a-Service) in einer privaten Cloud-Umgebung angeboten.

(3) Kanäle. Die Kanäle eines Geschäftsmodells beschreiben, wie das Unternehmen mit seinen Kunden in Kontakt tritt und ihnen das Wertangebot unterbreitet [25]. Der zentrale Vertriebskanal des AaaS-Anbieters ist das Internet. Der AaaS-Dienstleister wirbt darüber sowohl durch die eigene Webseite als auch über soziale Netzwerke für das Wertangebot. Darüber hinaus findet die Ansprache potenzieller Kunden auf Konferenzen und Prüfertagungen statt. Der Einsatz von Print-Medien ist grundsätzlich möglich, widerspricht allerdings der angestrebten Digitalisierungsstrategie.

(4) Kundenbeziehungen. Die Kundenbeziehungen beschreiben die Beziehungen, die mit dem Zielkundensegment eingegangen werden [25]. Dabei werden dem Unternehmen die Kundengewinnung, die Kundenpflege und die Verkaufssteigerung als Motivation unterstellt. Zu Beginn umfasst die Beziehung zu den Kunden, d. h. WPG und Mandanten, die Unterstützung bei der Einrichtung der Extraktionsroutinen. Langfristiger Bestandteil der Kundenbeziehungen von AaaS ist der on-demand self-service, der es den Kunden ermöglicht den Umfang und die Bestandteile des Wertangebots zusammenzustellen, ohne im direkten Kontakt mit dem Informationsdienstleister zu stehen. Zusätzlich werden jedem Kunden im Bedarfsfall eine persönliche Betreuung zugesichert. Als Kommunikationsinstrument dienen in diesem Zusammenhang auch die Service Level Agreements (SLAs), in denen der Umfang und die Leistung definiert sind.

(5) Umsatzmodell. Das Umsatzmodell beschreibt, wie dem Unternehmen die monetären Mittel durch die Bereitstellung des Wertangebots zufließen [25]. Die Nutzung der kontinuierlichen Prüfungsleistungen wird über eine Nutzungsgebühr abgerechnet. Die Höhe der Gebühr ergibt sich durch die tatsächlich genutzten Leistungen (engl. measured service). Eine Abrechnung entsprechend des für das CC typische pay-per-use-Prinzip kann z. B. pro Nutzer und Monat erfolgen. Darüber hinaus wäre die Abrechnung in periodischen Raten als Flatrate möglich, was den Nutzern einen unbegrenzten Zugang zu allen Funktionen der Plattform ermöglicht.

(6) Schlüsselressourcen. Die Schlüsselressourcen beschreiben die Fähigkeit, einen wiederkehrenden Prozess auf Basis interner und externer Ressourcen auszuführen [25]. Der Betrieb von AaaS erfordert ein Data-Center, in dem die IT-Infrastruktur, u. a. Hard- und Software, betrieben wird. Menschliche Expertise wird in den Bereichen der Softwareentwicklung, Marketing und Vertrieb sowie Wirtschaftsprüfung und IT-Revision benötigt.

(7) Schlüsselaktivitäten. Die Schlüsselaktivitäten und die Wertkonfiguration beschreiben die wichtigsten Aktivitäten zur Wertschaffung [25]. Die Schlüsselaktivitäten bei AaaS umfassen drei Kategorien: Die operativen Schlüsselaktivitäten umfassen die Implementierung von Extraktionsroutinen in den Systemen der Mandanten und Schnittstellen an das Audit-Datawarehouse. Zusätzlich werden Auswertungsskripte entwickelt und implementiert, die über den Datenbestand im Datawarehouse laufen. Weitere Aktivitäten betreffen Aufbau und Verwaltung der Infrastruktur sowie die Weiterentwicklung der eingesetzten Software. Spezielle Problemlösungsaktivitäten dienen zur Lösung spezifischer Kundenprobleme. Personalwirtschaftliche Aktivitäten sowie Marketing und Vertrieb deinen der Unterstützung des Betriebszwecks.

(8) Schlüsselpartner. Die Schlüsselpartnerschaften werden mit mindestens einem weiteren Unternehmen eingegangen, um den eigenen Wertschöpfungsprozess durch externe Ressourcen und Fähigkeiten zu ergänzen [25]. Für AaaS besteht die besondere Situation, dass die Kunden Wirtschaftsprüfer und Mandanten, auch als Partner fungieren, da sie Auswertungsregeln bzw. die Rechnungslegungsdaten liefern, die für die Erstellung des Wertangebots unerlässlich sind. AaaS kann daher als eine besondere Form des Co-Creation betrachtet werden. Weitere Schlüsselpartner sind andere Anbieter von Cloud-Services, von denen kurzfristig Speicherressourcen bezogen werden können, sowie externe Softwareentwickler. Darüber hinaus sind berufsständische Organisationen und Branchenverbände weitere Schlüsselpartner.

(9) Kostenstruktur. Die Kostenstruktur beinhaltet alle Kosten, die im Unternehmen durch die Erzeugung des Wertversprechens entstehen [25]. Kosten entstehen zum einen durch die Anschaffung und den Betrieb der technischen Infrastruktur, insb. durch ein Data-Center, die dem die Hardware betrieben wird. Zum anderen entstehen Kosten durch die Anstellung von Beschäftigten in den Bereichen Softwareentwicklung, Marketing und Vertrieb sowie der persönlichen Kundenbetreuung. Variable Kosten können zudem durch die Nutzung externer Ressourcen (engl. third-party-services), die sich ebenfalls nach dem pay-per-use-Modell berechnen, entstehen. Je mehr Kunden der AaaS-Dienstleister hat, desto mehr Infrastruktur muss extern bezogen werden. Weitere Kosten für Telekommunikation, elektrische Versorgung und Büroräume werden hier nicht weiterbetrachtet.

(8) Schlüsselpartner	**(7) Schlüsselaktivitäten**	**(2) Wertversprechen**	**(4) Kundenbeziehungen**	**(1) Kundensegmente**
▪ WPG: Liefern Auswertungsregeln Mandanten : Liefern Daten ▪ Cloud-Service-Provider ▪ Berufsständische Organisationen ▪ Branchenverbände	▪ Implementierung von Extraktionsroutinen ▪ Schnittstellen zu Data-Center/Audit-Datawarehouse ▪ Auswertungsskripte ▪ Aufbau/Veraltung Infrastruktur ▪ Personal, Marketing, Vertrieb	▪ Befähigung zur kontinuierlichen Prüfung von Rechnungslegungs-prozessen ▪ Jederzeit Zugriff auf Prüfresultate ▪ Zeitverzug der Prüfung verringern ▪ Belastungsspitzen reduzieren ▪ Datenqualität erhöhen ▪ Befreiung eigener, kapitalintensiver Infrastruktur	▪ Unterstützung bei Implementierung von Extraktionsroutinen ▪ on-demand self-service auf Plattform ▪ Persönliche Betreuung ▪ SLA	▪ Organisationen, die Rechnungslegung betreiben ▪ Organisationen, die Rechnungslegung prüfen ▪ WPG, kleine und mittlere WPG ▪ Interne Revision, Controlling, Business Intelligence von Großunternehmen ▪ Start-Ups ▪ Öffentlicher Sektor
	(6) Schlüsselressourcen ▪ Infrastruktur: Data-Center ▪ Menschliche Expertise ▪ Softwareentwicklung ▪ Marketing, Vertrieb ▪ Wirtschaftsprüfung/ IT-Revision		**(3) Kanäle** ▪ Internet (Webseite, soziale Netzwerke) ▪ Vor Ort (Prüfertagungen, Messen)	

(9) Kostenstruktur	**(5) Umsatzmodell**
▪ Technische Infrastruktur ▪ Softwareentwicklung ▪ Personal, Marketing, Vertrieb ▪ Persönliche Kundenbetreuung	▪ Nutzungsgebühr über measured services ▪ Pay-per-use-Prinzip: Abrechnung pro Nutzer pro Monat ▪ Periodische Raten: Flatrate

Abbildung 2. Audit-as-a-Service Business Model Canvas

4.3 Geschäftsmodellanalyse

Technologischer Fortschritt, insbesondere das als „Basisinnovation“ [39] charakterisierte CC, verändert die Geschäftsmodelle von IT-Dienstleistern [1]. Daraus ergibt sich die Notwendigkeit, Geschäftsmodelle hinsichtlich der cloud-charakteristischen Umsetzbarkeit zu beurteilen. Labes et al. (2014) vereinen dazu bestehende Geschäftsmodellelemente und cloud-spezifische Gestaltungsmerkmale zu einem morphologischen Kasten, der als Ordnungsrahmen für die Analyse cloud-basierter Geschäftsmodelle genutzt werden kann [16]. Das im vorherigen Abschnitt vorgestellte, auf einer Cloud-Architektur basierende Geschäftsmodell für die kontinuierliche Prüfung wurde anhand dieses Ordnungsrahmens analysiert (siehe Abbildung 3). Wie nachfolgend erläutert, konnten dabei kritische Elemente identifiziert werden, die bei der Anwendung des Business Models Canvas bislang nicht berücksichtigt wurden.

Strategie. Bei der Erstellung des Geschäftsmodells wurden bislang keine strategischen Aspekte betrachtet. Mit AaaS wird vorrangig eine Differenzierungs-strategie angestrebt, da sich der Betriebszweck, nämlich die kontinuierliche, unterjährige Prüfung von Rechnungslegungsprozessen, vom traditionellen, jährlichen Vorgehen anderer Marktteilnehmer maßgeblich unterscheidet. Zu Teil ist die Strategie als Nischenstrategie zu begreifen, weil AaaS hauptsächlich kleine und mittlere WPG unterstützt. Durch die Gestaltung (engl. design) eines innovativen Prüfungsansatzes kann AaaS als Markterweiterung betrachtet werden. Durch die unternehmensübergreifende Zusammenführung von Daten des Mandanten und Auswertungsregeln der WPG wird letztendlich auch eine Marktdurchdringung (engl. diffusion) im Sinne einer Steuerung von Akteuren angestrebt. Als Bindeglied zwischen WPG und Mandanten fokussiert Audit-as-Service eine vertikale Integration.

Wertversprechen. Die bei der Geschäftsmodellerstellung angestellten Überlegungen zum Wertversprechen konnten durch die Analyse erweitert werden. Die cloud-spezifischen Leistungen sind durch die Bereitstellung eines Audit-Datawarehouses als zentralen Speicherort für die Rechnungslegungsdaten gekennzeichnet. In dem Audit-Datawarehouse werden Auswertungsregeln ausgeführt (engl. computing). Zusätzlich werden den Kunden eine Entwicklungsumgebung mit eigenen Werkzeugen angeboten, mit denen individuelle Auswertungsregeln erstellt werden können. Gesetzliche Vorgaben an die Auslagerung von Rechnungslegungsdaten erfordern den Betrieb einer privaten Cloud-Infrastruktur, bei der der physische Speicherort der Daten jederzeit bestimmt werden kann. Dadurch ist der Standort der Leistung beschränkt ist. Darüber hinaus skaliert der Leistungsumfang vom Speichervolumen der Rechnungslegungsdaten und den gewählten Auswertungsregeln, die durch den Benutzer individualisiert werden können.

Wert generieren. Durch die Analyse des Geschäftsmodells konnten weitere Entwicklungspotenziale in Bezug auf die Generierung des Werts identifiziert werden. Ziel des Ansatzes ist die Errichtung eines strategischen Netzwerks, das die Unterstützung weiterer Partner mit fachlichem und technischem Wissen benötigt. Neben umfangreichen technischen Ressourcen müssen dazu auch Aktivitäten zum Aufbau und Verwaltung einer sachgerechten Infrastruktur durchgeführt werden. Eine besondere Schwierigkeit stellt die Abschätzung der Kostenstruktur dar. Neben klassischen, variablen Komponenten, wie z. B. Gehältern, könnten durch eine Überauslastung der Systeme kurzfristig weitere Speicherkapazitäten notwendig werden. Dadurch müssten Speicherkapazitäten bei externen Technologiepartnern eingekauft werden, womit erhebliche variable Kosten verbunden wären.

Wert vertreiben. Das Geschäftsmodell sieht eine variable Generierung von Erlösen als pay-per-use-Modell vor. Ebenso ist jedoch auch die Buchung eines Basispakets denkbar, das über periodische Zahlungen, entsprechend der eingestellten Prüfungsintervalle, abgerechnet wird. Wie im Geschäftsmodell beschrieben, ist der maßgebliche Vertriebsweg von AaaS eine Netzwerk-Infrastruktur, auf die Benutzer über das Internet zugreifen können. Um das Wertversprechen einer ortsunabhängigen Prüfung sicherzustellen, muss eine Darstellung von Ergebnissen auf mobilen Endgeräten unterstützt werden. Darüber hinaus erfordert die Implementierung von Extraktionsschnittstellen die Erbringung von Services beim Mandanten vor Ort. AaaS kann vorwiegend als Unterstützungsleistung (eng. support) betrachtet werden. Da es sich bei Rechnungslegungsdaten um sensible, hochkritische Daten handelt, die durch gesetzliche Vorgaben geschützt werden, müssen Informationen über deren Auswertung und Speicherung über transparente SLAs kommuniziert werden. Der Zielmarkt wurde durch das Business Model Canvas ausreichend adressiert. Der Markt- und Kundenfokus liegen auf einer klar abgrenzbaren Branche (Wirtschaftsprüfung) und zum Teil in der Nische der kleinen und mittleren WPG sowie internen Revisionseinheiten von Großunternehmen.

	Kategorie	Unterkategorie	Gestaltungsmerkmale						
	Strategie	Generische Strategie	Kostenführerschaft	Differenzierungsstrategie	Nischenstrategie				
		Marktstrategie	Market Adaption	Market Design	Market Diffusion	Market Co-construction			
		Markteintritt	Neueintritt	Markterweiterung	Know-how-Transfer	Vorherige Markterfahrung			
		Wertschöpfung	Horizontal	Vertikal					
	Wertversprechen	Leistung („as-a-Service“)	Speicher	Computing	Netzwerk	Entwicklungs-umgebung	Entwicklungs-werkzeug	Software	Geschäfts-prozesse
		Bereitstellungsmodell	Private	Community	Hybrid	Public			
		Service-Typ	Angebot	Aggregation	Aggregation mit Zusatz	Vergleich und Kategorisierung	Integration	Beratung	
		Eigenschaften	Skalierbarkeit	Individualisierbarkeit	Standortbeschränkung	Interoperabilität			
Wert generieren	Partner-Netzwerk	Netzwerkart	Ecosystem	Strategisch	Lose	Keine			
		Partnerart	Technologie	Business	Consulting				
		Geschäftsfeld	Fremdes Geschäftsfeld	Ähnliches Geschäftsfeld	Gleiches Geschäftsfeld				
	Ressourcen & Tätigkeiten	Ressourcen	Hardware	Software	Netzwerk	Daten/Inhalte	Know-how	Personal	
		Aktivitäten	Infrastruktur-verwaltung	Personal-wirtschaft	Entwicklung	Beschaffung	Eingangs-logistik	Ausgangs-logistik	Marketing
	Kosten	Abhängigkeit von Ausbringungsmenge	Hauptsächlich Fixkosten	Hauptsächlich variable Kosten					
Wert vertreiben	Erlöse	Nutzer-Zahlungsmodell	Einmal-gebühren	Periodische Raten	Reservierung	Pay-per-use	Spot	Kostenfrei	
		Partner-Zahlungsmodell	Sponsoring	Werbung	Umsatzbeteiligung				
	Vertrieb und Kunden-beziehung	Kanal	Internet	Mobil	Print-Medien	Vor Ort			
		Kundenbeziehung	Selbstservice	Online-Profil	Community	Support	Transparente SLAs		
	Zielmarkt	Marktfokus	Masse	Branche	Nische				
		Kundenfokus	Großunternehmen	KMU	Start-Ups	Öffentlicher Sektor	Verbraucher		

Abbildung 3. Morphologischer Kasten von Audit-as-a-Service in Anlehnung an [16]

5 Kritische Würdigung

Ausgangspunkt der durchgeführten Untersuchung war die mangelnde Verbreitung von kontinuierlichen Prüfungsansätzen in der Praxis. Wesentliche Ursachen dafür sind der umfangreiche Personal- und Wissensbedarf sowie zahlreiche technische und organisatorische Barrieren, die vor allem kleine und mittlere WPG an einer Umsetzung kontinuierlicher Prüfung hindert. Ein Lösungsansatz ist die Auslagerung der Aktivitäten zur Umsetzung und Durchführung der kontinuierlichen Prüfung an einen spezialisierten Informationsdienstleister. In Hinblick auf die eingangs formulierte Forschungsfrage wurde in diesem Artikel die Repräsentation eines auf einer cloud-Architektur basierenden, innovativen Geschäftsmodell in der Wirtschaftsprüfung entworfen. Dazu wurde das Business Model Canvas nach Osterwalder (2004) und Osterwalder und Pigneur (2010) verwendet.

Das vorgestellte Geschäftsmodell ist lediglich als erster, aber notwendiger Entwurf zu verstehen. Je nach Ausgestaltung der einzelnen Bausteine können unterschiedliche Geschäftsmodelle generiert werden. Es ist festzuhalten, dass es keine einzig richtige Lösung für ein Geschäftsmodell zu Audit-as-a-Service geben kann, sondern verschiedene, nützliche Geschäftsmodelle in Abhängigkeit von den beteiligten Parteien und der angebotenen Leistung möglich sind. Wir gehen davon aus, dass eine besonders nützliche Lösung, d.h. ein sehr geeignetes Geschäftsmodell, erst näherungsweise durch mehrere Evaluationszyklen und schließlich erst bei einer Umsetzung unter realistischen Bedingungen entstehen wird.

Ein zweiter, wesentlicher Kritikpunkt ist, dass vermutlich nicht alle technischen und organisatorischen Barrieren durch die die Auslagerung und das entworfene Geschäftsmodell überwunden werden können. Vielmehr findet eine Verlagerung der Umsetzungsrisiken auf den spezialisierten Informationsdienstleister statt. Dem kann entgegengehalten werden, dass sich der Informationsdienstleister das Risiko durch die gezahlten Tantiemen entschädigen lassen und im Laufe der Zeit durch die Generierung von Erfahrungskurveneffekte die Leistung günstiger anbieten können wird.

Aus methodischer Sicht kann der Vergleich von lediglich sieben Konzepten und die Anwendung eines einzigen Ansatzes kritisch betrachtet werden. Eine weitere Limitation dieses Artikels kann daher in den theoretischen Ausführungen zur Auswahl und Begründung der betrachteten Geschäftsmodellkonzepte gesehen werden. Allerdings konnte festgestellt werden, dass wesentliche inhaltliche Überschneidungen zwischen den betrachteten Geschäftsmodellkonzepten bestehen.

Eine weitere Limitation liegt in den Ausführungen zu den Gestaltungsmerkmalen des CC im Business Model Canvas. In der Literatur werden für das CC zahlreiche Aspekte diskutiert, die im Rahmen dieses Artikels nicht betrachtet wurden. Durch Vorstudien konnte jedoch festgestellt werden, dass die Kerneigenschaften in der Literatur sehr ähnlich beschrieben und zum Teil standardisiert sind [40].

6 Fazit und Ausblick

Zusammengefasst kann das vorgestellte Geschäftsmodell als erste, konzeptionelle Repräsentation einer neuartigen Geschäftslogik verstanden werden, die durch einen iterativen Prozess verbessert werden muss. Damit werden die bislang vorrangig technologischen Betrachtungen der digitalen Transformation der Abschlussprüfung um Aspekte der Geschäftsmodellentwicklung erweitert. Zudem werden mit dem Geschäftsmodell erstmalig die Möglichkeiten aufgezeigt, wie Eigenschaften die Cloud-Computings für die Umsetzung der kontinuierlichen Prüfung genutzt werden können. Der Beitragt zeigt somit, wie IKT das bestehende Geschäftsmodell der traditionellen Jahresabschlussprüfung verändert und zu einem neuen, disruptiven Ansatz weiterentwickelt. Dadurch können analoge Vorgehensweisen für die digitale Transformation anderer klassischer Dienstleistungen abgeleitet werden.

Die Autoren dieses Artikels erwarten mit den vorgestellten Ergebnissen neue Impulse für die digitale Transformation der Abschlussprüfung und die, insbesondere im deutschsprachigen Raum, stagnierte Debatte über die kontinuierliche Prüfung geben zu können. Gegenstand geplanter Forschungsarbeiten ist die Evaluation des erstellten Geschäftsmodells durch Experten aus der IKT-Branche und die prototypische Umsetzung der Kernkomponenten um auf dieser Basis parallele Simulationen zur traditionellen Abschlussprüfung durchzuführen. Zukünftige Arbeiten sollten zudem die Preisbereitschaft für die Nutzung von AaaS über einen quantitativen Querschnitt der Kundensegmente untersuchen. Fernziel ist die Ausgründung einer Gesellschaft, die die wirtschaftliche Verwertung der vorgestellten Geschäftsmodellinnovation anstrebt.

Referenzen

1. Veit, D., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Kundisch, D., Leimeister, J.M., Loos, P., Spann, M.: Business models. Bus. Inf. Syst. Eng. 6, 45–53 (2014).
2. Rega, I., Teipel, G.: Digitalisierung in der Wirtschaft und im Berufsstand. Die Wirtschaftsprüfung. 69, 39–45 (2016).
3. Wilting, A.: Braucht INDU5TR1E 4.0 den W1RT5CH4FT5PRÜFER 2.0? Die Wirtschaftsprüfung. 14, (2014).
4. Marten, K.-U., Quick, R., Ruhnke, K.: Wirtschaftsprüfung. Schäffer-Poeschel Verlag, Stuttgart (2015).
5. Rezaee, Z., Sharbatoghlie, A., Elam, R., McMickle, P.L.: Continuous auditing: Building automated auditing capability. Audit. A J. Pract. Theory. 21, 147–163 (2002).
6. Leimeister, J.M.: Einführung in die Wirtschaftsinformatik. Springer-Verlag Berlin Heidelberg (2015).
7. Alles, M.G., Kogan, A., Vasarhelyi, M.A.: Putting continuous auditing theory into practice: Lessons from two pilot implementations. J. Inf. Syst. 22, 195–214 (2008).
8. Singh, K., Best, P.J., Bojilov, M., Blunt, C.: Continuous Auditing and Continuous Monitoring in ERP Environments: Case Studies of Application Implementations. J. Inf. Syst. 28, 287–310 (2014).
9. Byrnes, P.E., Ames, B., Vasarhelyi, M., Warren Jr., J.D.: AICPA: The Current State of Continuous Auditing and Continuous Monitoring (Whitepaper). New York, USA (2012).
10. Kiesow, A., Zarvić, N., Thomas, O.: Improving the Success of Continuous Auditing Projects with a Comprehensive Implementation Framework. In: Proceedings of the European Conference on Information Systems (ECIS 2015). AIS, Münster, Germany (2015).
11. Kempf, D.: Aktuelle Entwicklungen und Trends in der IT. Die Wirtschaftsprüfung. 66, (2013).
12. Debreceny, R.S., Gray, G.L., Ng, J.J.-J., Lee, K.S.-P., Yau, W.-F.: Embedded audit modules in enterprise resource planning systems: implementation and functionality. J. Inf. Syst. 19, 7–27 (2005).
13. Kiesow, A., Thomas, O.: Digitale Transformation der Abschlussprüfung. Die Wirtschaftsprüfung. 69, 709–716 (2016).
14. Wöhe, G., Döring, U.: Einführung in die Allgemeine Betriebswirtschaftslehre. S. 166f. Verlag Franz Vahlen. München (2013).
15. Vasarhelyi, M.A., Alles, M.G.: The “now” economy and the traditional accounting reporting model: Opportunities and challenges for AIS research. Int. J. Account. Inf. Syst. 9, 227–239 (2008).
16. Labes, S., Hahn, C., Erek, I.K., Zarnekow, R.: Geschäftsmodelle im Cloud Computing. In: Digitalisierung und Innovation. pp. 35–60. Springer Gabler, Wiesbaden 2013 (2013).
17. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. Manag. Inf. Syst. Q. 26, 3 (2002).
18. Kiesow, A., Schomaker, T., Thomas, O.: Transferrring Continuous Auditing to the Digital Age - The Knowledge Base after three Decades of Research. In: Proceedings of the European Conference on Information Systems (ECIS 2016). Istanbul, Turkey (2016).
19. Kneer, D.C.: Continuous Assurance: We are Way Overdue. Inf. Syst. Control J. 1, 30–34 (2003).
20. Greenstein, M.M., Ray, A.W.: Holistic, continuous assurance integration: e-business opportunities and challenges. J. Inf. Syst. 16, 1–20 (2002).
21. Murthy, U.S., Groomer, S.M.: A continuous auditing web services model for XML-based accounting systems. Int. J. Account. Inf. Syst. 5, 139–163 (2004).

22. Yeh, C.-H., Shen, W.-C.: Using continuous auditing life cycle management to ensure continuous assurance. African J. Bus. Manag. 4, 2554–2570 (2010).
23. Schallmo, D.: Geschäftsmodell-Innovation - Grundlagen, bestehende Ansätze, methodisches Vorgehen und B2B-Geschäftsmodelle. Springer-Verlag, Wiesbaden (2013).
24. Morris, M., Schindehutte, M., Allen, J.: The entrepreneur 's business model: toward a unified perspective. J. Bus. Res. 58, 726–735 (2005).
25. Osterwalder, A.: The business model ontology: A proposition in a design science approach. Inst. d'Informatique Organ. Lausanne, Switzerland, Univ. Lausanne, Ec. des Hautes Etudes Commer. HEC. 173, (2004).
26. Amit, R., Zott, C.: Value creation in e-business. Strateg. Manag. J. 22, 493–520 (2001).
27. Bieger, T., Reinhold, S.: Das Wertbasierte Geschäftsmodell - Ein aktualisierter Strukturierungsansatz. In: Innovative Geschäftsmodelle - Konzeptionelle Grundlagen, Gestaltungsfelder und unternehmerische Praxis. pp. 13–70 (2011).
28. Chesbrough, H., Rosenbloom, R.S.: The role of the business model in capturing value from innovation: evidence from Xerox Corporation ' s technology spin-off companies. Ind. Corp. Chang. 11, 529–555 (2002).
29. Labes, S., Zarnekow, R.: Geschäftsmodelle im Cloud Computing. In: Wirtschaftsinformatik in Wissenschaft und Praxis. pp. 179–190. Springer Berlin Heidelberg (2014).
30. Schallmo, D.: Bestehende Ansätze zur Business Modell Innovation - Analyse und Vergleich der Geschäftsmodelle. Wiesbaden (2015).
31. Bieger, T., Reinhold, S.: Das wertbasierte Geschäftsmodell–Ein aktualisierter Strukturierungsansatz. In: Innovative Geschäftsmodelle. pp. 13–70. Springer (2011).
32. Magretta, J.: Why business models matter. Harv. Bus. Rev. 80, 86–92 (2002).
33. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Q. 28, 75–105 (2004).
34. Thomas, O.: Management von Referenzmodellen: Entwurf und Realisierung eines Informationssystems zur Entwicklung und Anwendung von Referenzmodellen. Logos, Berlin (2006).
35. Kiesow, A., Thomas, O.: Continuous Auditing Systeme: Rahmenwerk zur Gestaltung von Informationssystemen für kontinuierliche Prüfungsdienstleistungen. In: Multikonferenz Wirtschaftsinformatik (MKWI). Ilmenau, Germany (2016).
36. BITKOM: Wie Cloud Computing neue Geschäftsmodelle ermöglicht. Berlin (2013).
37. Kiesow, A., Schomaker, T., Thomas, O.: Konstruktion von Prozessmodellen für digitalisierte Prüfungsdienstleistungen. In: Smart Systems Engineering (2016).
38. Osterwalder, A., Pigneur, Y.: Business model generation: a handbook for visionaries, game changers, and challengers. John Wiley & Sons, Hoboken, New Jersey (2010).
39. Repschläger, J., Pannicke, D., Zarnekow, R.: Cloud Computing: Definitionen, Geschäftsmodelle und Entwicklungspotenziale. HMD Prax. der Wirtschaftsinformatik. 47, 6–15 (2010).
40. Mell, P., Grance, T.: The NIST Definition of Cloud Computing - Recommendations of the National Institute of Standards and Technology. (2011).

Fostering Business Model Extensions for ICT-Enabled Human-Centered Service Systems

Stefan Kleinschmidt[1], Christoph Peters[1,2]

[1] University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland
{stefan.kleinschmidt,christoph.peters}@unisg.ch
[2] University of Kassel, Research Center for IS Design (ITeG), Kassel, Germany
christoph.peters@uni-kassel.de

Abstract. When improving human-centered service systems (HCSSs) with information and communications technology (ICT), financial aspects are important but challenging for companies with established business models (BMs). The use of ICT and changes in value creation reflect business needs, but commercial success requires modifications and extensions of the BMs. However, prevailing approaches do not take account of these requirements. In this paper, we present a BM design process that fosters the extension of BMs for ICT-enabled HCSSs to support service innovations. Using an action research project in the field of volunteering, we iterated and revised the BM design process in a project collaboration with three end-user companies having similar objectives. The process guides those responsible for service innovation in structuring, analyzing, and the decision-making of alternative BM extensions. Thus, the presented approach contributes to ICT-related service innovation projects by describing systematic and repeatable activities that are the first step for commercial success.

Keywords: Business Model Design, Service Design, Service Innovation, ICT, Action Research

1 Introduction

From a business perspective, leveraging information and communications technology (ICT) to improve human-centered service systems (HCSSs) makes sense as it enables productivity improvements, cost reductions, and innovation [1, 2]. Visionary companies take advantage of this and put established competitors under pressure to respond to new trends and rethink their value creation [3, 4]. While it appears that not all of the ICT-enabled service innovations solely focus on the strategic competitive advantages, the commercial success of the innovations is important [5]. Here, ICT has to generate value in the corresponding business model (BM) [1]. While companies have problems to generate successful and sustainable BMs [6], neither general approaches that foster BM extensions for ICT-enabled HCSSs exist [7] nor do existing BM approaches cover the required activities of BM design [8].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Kleinschmidt, S.; Peters, C. (2017): Fostering Business Model Extensions for ICT-Enabled Human-Centered Service Systems, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 897-911

Both research and business practices have recognized the potential of BMs for ICT-enabled HCSSs. In research, the influence of ICT on service innovations and the development of corresponding BMs are named as top research priorities [9, 10]. The connection between the two concepts is seen as important because digital BMs can change the way value is created in ICT-enabled services [7, 11]. Moreover, design knowledge for BMs includes an explicit contribution answering these strategic business questions [12]. In business practice, the commercialization of ICT-enabled HCSSs can have profound economic consequences within an emerging market [2].

Given the importance of the systematic extension of BMs for ICT-enabled service innovations in HCSSs, it is surprising that little research has been done on this topic. Therefore, the research question that is to be answered in this paper is:

How can business models be adapted to commercialize ICT-enabled service innovations in human-centered service systems?

We use an action research project in the field of volunteering to derive a systematic and repeatable BM design process. This process can foster the extension of BMs for supporting ICT-enabled service innovations in HCSSs and facilitating their commercial success. We derive the process based on the existing literature on BM design and test and iterate it in three different project settings. We use the learnings from each case to improve the BM design process that aims at structuring, analyzing, and the decision-making of alternative BM extensions. In conclusion, this paper contributes to the existing literature with an explicit and systematic approach to BM design in HCSSs that considers previous theoretical findings and solves real-world problems.

The paper is structured as follows. After this introduction, the next section provides the theoretical backgrounds of ICT-enabled HCSSs and BM design, which define the key terms and concepts relevant to these related fields. The third section justifies and describes the action research approach used in this paper. The fourth and fifth section contain the derivation of the BM design process from the literature and learnings from the three use cases in our project. The application of this process is described in the sixth section before we then discuss it with regard to its improvements and limitations in the seventh section. In the end, the findings are summarized and a proposal for further work is presented.

2 Theoretical Background

2.1 ICT-Enabled Human-Centered Service Systems

Service and service systems are essential concepts for the understanding of HCSSs. In line with the service-dominant logic, we understand service as a process cocreating context-specific value [13]. It takes place in service systems as it includes several participants, with one of them being the beneficiary. These service systems are "configurations of people, information, organizations, and technologies that operate together for mutual benefit" [2]. Put simply, services are problem-solving based on the capabilities and interaction between the different actors. The various parts are connected via coordination and cooperation to create services [13].

HCSSs, which are service systems with a focus on human interaction and personal services, require particular attention [2]. They differ from other service systems because personal interaction is essential for the value creation [14]. These properties are present in diverse industries such as hospitality, healthcare, and retail, but not limited to these [1, 2]. In this paper, we base our findings on the field of volunteering. Volunteers are part of HCSSs as volunteering can be defined as “giving time or skills during a planned activity for a volunteer group or organization” [15], which implies interaction and personal engagement.

Based on emerging trends such as an increase in digital innovation, there is an expansion of ICT in service systems [1]. ICT is fundamental to many HCSSs [4, 16]. Here, ICT and people are an integral part of the ICT-enabled value creation [17]. As coordination and cooperation are critical to value creation, the leveraging of ICT for HCSSs influences how HCSSs work. Furthermore, ICT enables service innovation that helps to improve and expand the service systems [1, 9]. On the one hand, ICT can enhance the efficiency of current services by simplifying coordination and cooperation. On the other hand, ICT offers possibilities to come up with new services when integrating customers and employees into the service process and substituting them [9].

The presented properties produce a number of challenges. Service innovation, a service system reconfiguration to increase the value for the involved actors [18], is important to remain successful [9]. This is difficult to implement as HCSSs resist traditional optimization solutions and their economic contribution is hard to measure [2]. Also, the dependency on knowledge and customization, as it is the case in volunteering, increases customer involvement in value creation [19]. The support of ICT could solve these challenges [17]. Here, ICT-enabled HCSSs have to prove their benefit over traditional service systems and require mechanisms to build synergies with existing systems [5]. For the systematic design of service systems, concepts of service engineering have been expanded to the system’s perspective, which are coined under the term service systems engineering [20]. So far, none of these concepts, methods, or design processes cover a BM perspective.

2.2 Business Model Design

BMs and their components cover the essential functions of service systems. A BM is a “simplified and aggregated representation of the relevant activities of a company. It describes how marketable information, products and/or services are generated by means of a company's value-added component” [10]. The objective of BM design is to expose business opportunities by defining a value for the participants involved in the service system [3]. The customer value proposition that “describes the benefits customers can expect from products and services” reflects this purpose [21]. Thus, it also shows how a client's problem is solved [22]. This approach is related to how a service is designed and can therefore not be considered separately [6, 23]. A BM refers mostly to a particular product or service with an associated value to the customer [21]. This value to the customer is the basis for the financial return of the product or service [22]. BM design is a critical task for managers and represents a challenge when established BMs are reconsidered and revised [3].

The result of BM design is a set of relevant and individual activities that represent the positioning of the company to customers and competitors. The process of the design is similar in many cases although there is no uniform approach [8]. What is challenging to BM design is that contrary to partial optimization, system-level design is important [3]. A BM that only focuses on value creation and/or delivery without simultaneously considering revenue streams from that value will likely be economically unsustainable [24]. Additionally, when reconsidering existing BMs, there are forces of inertia and resistance to change that represent a challenge [3].

The activities in BM design are supported by different concepts. These concepts represent interfaces to BM design [10]. They help in the development and commercialization of service systems and their corresponding BMs. Thus, we searched for research streams that follow these objectives. To verify the completeness of our findings, we asked two experts of the field for input on such concepts. Table 1 lists the concepts including their objectives and the connection they have with BM design.

Table 1. Concepts Related to Business Model Design

Concept	*Objective*	*Interfaces*
Customer Development [25]	Align product development with customer needs and market	Service ideation, revenue model
Lean Startup Approach [26]	Service development with shortened development cycles	Service development and implementation
Service Design [27]	Design service according to the needs of customers or participants	Service ideation and development
Service Innovation [1]	Reconfiguration to increase the value for the involved actors	Service ideation and development
Value Proposition Design [21]	Define the functionality of services and customer needs	Service ideation, development, and visualization

3 Research Method

Canonical action research combines the generation of scientific knowledge with researcher intervention to solve immediate real-world problems in a formalized form [28, 29]. Here, the fundamental assumption is that an understanding of complex processes can be achieved by changing these processes as well as observing and reflecting the effects. Good results can be expected when the goals of the researchers coincide with those of the field partners and the researchers are actively involved in the process of problem solving [28]. As action research runs in iterations, the cycles of diagnosing, action planning, action taking, evaluating, and specifying learning are repeated until the problem is solved [30]. IS research and HCSS settings such as education have successfully used this research approach in the past [29].

The research presented in this paper was embedded in a research project that aimed at developing and introducing online matchmaking platforms for volunteers. The platforms support organizations to coordinate the placement of volunteers and people or nonprofits in need of help. The research project contained a technical part including the development and testing of the matchmaking software and a business part including the design and implementation of the service system. The project collaboration consisted of researchers, technical developers, and the future operators of the platform. During the project, it became apparent that a BM design was needed based on existing services that had not been available in the prevailing approaches. The objective of the BM extension was to help those responsible for the service financing, improvement, and further operation.

The data collection was initiated in a joint problem analysis with all project partners. Based on the problem description, we conducted a literature review with keywords of "business model design", "customer development", lean startup", "service design", "service innovation", and "value proposition design". High-quality literature was searched for processes that guide the BM extensions for ICT-enabled service innovations in HCSSs. Qualitative data analysis [31] was applied to the found literature to search where the design process starts and ends and what the different steps in between are. Based on the results and with the method of collaboration engineering [32], we planned workshops in three organizational settings (Table 2). We had the chance to develop a BM for every company. Therefore, the three settings represented the research cycles in the action research project. The BM design process took place in seven workshops. The analysis was done in one workshop with all partners. In company A and C, the remaining process was completed in one workshop with a larger group of experts. In company B, the BM design was performed in four internal workshops. Based on the learnings from the previous workshops, we revised the BM design and reapplied it to the next setting. A documentation of the actions was made in the form of protocols.

Table 2. Organizational Settings (anonymized company profiles)

Company	A	B	C
Industry	Regional Development	Volunteering	Care Service
Country	Germany	Switzerland	Germany
Staff	15	75	940
Project Owner	General Manager	Product Manager	Project Manager
Participants	14 (8 external experts)	4	9 (5 external experts)

4 Knowledge Base

4.1 Literature Review

The BM design process is intended to help those responsible for the HCSSs in financing their existing and ICT-enabled service processes by guiding the extension of the BMs. Therefore, the process steps should focus on design including commercialization. We found ten publications that specifically deal with the design and extension of service systems and BMs. Table 3 includes the findings regarding the found publications and their coverage of the phases in the order of their publication date.

Table 3. Results of the Literature Review

Publication	*Steps of the Business Model Design Process*			
	Analysis	Idea Generation	Evaluation & Selection	Implementation
[33]	○	○	◐	○
[3]	○	●	◐	○
[23]	●	●	◐	●
[24]	●	●	◐	◐
[34]	○	●	●	◐
[35]	○	●	○	●
[36]	●	●	○	●
[37]	◐	●	○	○
[8]	◐	●	○	●
[38]	●	●	○	●
● = Covered ◐ = Partially Covered ○ = Not Covered				

Different steps in the process that could be identified include the analysis, idea generation, evaluation and selection, and implementation. None of the publications covers the full BM design process. First, the analysis is mostly a preparatory activity that enables an effective and efficient design of the BM extensions. Because the BM is to be built on an existing service enabled by ICT, the analysis should include a particular business service. A previous analysis of the environment is therefore not necessary. The result could include a visualization of the current service design and BM including ICT [23]. This is used to challenge the individual activities of the service [8]. Second, the idea generation contains the conception of the BM extension. This is most of the time the essential part of the design process. Here, the possible BMs are developed in terms of the definition of activities that generate value for a particular customer segment, specify the service provider, and link the activities with each other [3]. The objective is to conceptualize value creation for the customer and derive revenues for the service provider. The result could be a collection of possible and realistic ideas [23]. Nearly all publications cover this step. Third, the evaluation and selection aim at assessing the developed BM extension and choosing the most promising option(s). This could include the critical questioning of the BM design decisions and the financial contribution [34]. The evaluation outcome depends on the impact on the customer value, the sustainability of the investment and innovation, and the ability to compete with competitors [33]. The result could be a clear idea of which options are available, particularly for quick and long-term successes, and a selection of the most promising BM extension [35]. Fourth, the implementation builds a starting point for a functioning BM as it integrates the transition from concept to operationalization. Therefore, the necessary steps for an implementation are defined including the associated risks [24]. The initiation of quick wins is an integral part of this action [23]. After that, the implemented BM is influenced by a revision and reconfiguration based on external feedback and learnings [8]. This revision ensures a design according to the customer needs.

In summary, the literature reviews show that none of the publications depicts the complete BM design process. Findings were derived with regard to the steps of the BM design process. An explicit application to the service system that focuses on human interaction and personal services is not included. Therefore, this adaptation is part of the workshop preparation and follow-up.

4.2 Workshop Insights

Table 4. Results from the Workshops

Company	A	B	C
Diagnosing	Needs more users so that it is profitable.	Needs further use cases for continuous funding.	Introduced and needs cross-financing.
Action Planning	Definition of possible users of the platform.	Definition of extended users and use.	Definition of additional use with revenue.
Action Taking	Focus on the development of BMs via exploitation of existing resources.	Focus on BMs with new customers through new use cases.	Focus on the definition of customer groups with associated values and revenues.
Evaluating	BM extensions were developed for other age groups and a supply service.	BM extensions were developed for municipalities and companies including implementation steps.	BM extensions were developed for other service providers, local care facilities, and city authorities including a financial concept.
Specifying Learning	More focus on concrete customer values with associated revenues.	Further steps are necessary to implement the design.	Extensions including customer groups, value propositions, and revenue streams.

Following the action research cycles introduced in the method section, the different stages to improve and iterate the findings were completed. Table 4 includes the results from the workshops. The results from the diagnosing phase show that the objective of the BM extension is similar in all cases of the project. The primary problem throughout the cases was that the service was improved by ICT to support the offline service system with a value proposition that remains constant. As this is a platform business and primary customers are not willing to pay (they are volunteers), this challenge included the expansion of the use of ICT-enabled service systems to add additional customers who cover the costs. On this basis, the action planning and action taking for each workshop that describe the purpose of the workshop sessions were defined. The objective was to find new users, new uses, and new revenue opportunities. The learnings from the previous workshops initiated the changes in the workshop. While the first workshop focused on the entire BM with its activities, the second focused on specific customer groups, which was more effective. Based on these customer groups, the associated value proposition, and corresponding revenue opportunities were defined. The learning from the second case was that the design includes the definition of implementation steps. The evaluation of the BM extensions was done in the form of a workup of the workshop results. The experts who participated in the workshops commented on the quality of the BMs. In all companies, BM extensions could be triggered successfully.

5 Result

The result of the iteration and revision is the BM design process depicted in Figure 1. The process includes all activities that are necessary to create BM extensions for ICT-enabled HCSSs. The first three steps, depicted in gray, include activities that are directly related to the design process. The activities in the fourth step, the implementation, affect the BM design retrospectively and guarantee the desired service and BM design.

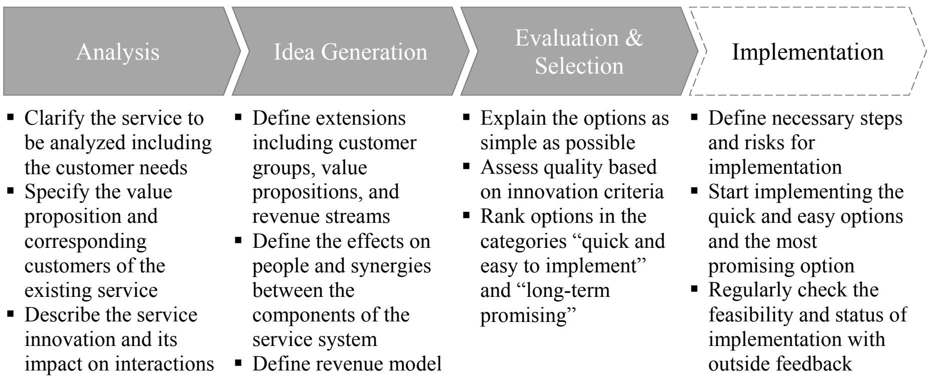


Figure 1. Business Model Design Process

The first step is the analysis of the existing service. Here, the service and the BM that is to be extended are identified and clarified. The value of the service to stakeholders and the activities in the form of a value proposition are an important part thereof [36]. The value proposition canvas [21] can be used to describe the service and specify the customer needs and corresponding values. Moreover, for a shared understanding of the purpose and service output, the effects of innovations from ICT are described. This provides the opportunity to show all results, constraints, and interrelationships of the service system [3]. The analysis challenges the individual activities of the service system [8] with a particular focus on the customers and their value proposition [21]. In HCSSs, the identification of capabilities and interaction is important as they constitute the value cocreation [2]. For a description of all other activities included in the BM and service innovation, the business model canvas [23] can be used.

The second and most important step in BM design is the idea generation. Here, the BM extensions are defined in a structured manner including a customer segment, a matching value proposition, and the value capture expected for the service provider

[24]. There are different starting points for the collection of ideas [23, 24]. The basis for the considerations of BM extensions in HCSSs are the ICT-enabled service innovations that include some new BMs [17]. Desirable are innovations originating from an enhanced customer value that put the customer at the center of service innovation and innovations where the technology is the driver for the BM extension [35]. Here, it is important to know which parts are visible and desirable to the customer [24]. A funding model complements a BM design [3, 23]. It is the result of the division of investments, costs, and revenues. In the end, the options must be assessable regarding the effects and synergies [34].

The third step is the evaluation and selection of alternative BM extensions. Here, the options that have been developed in the previous step are presented to increase a shared understanding. Based on this outline, the quality of the extensions is assessed regarding the BM design decisions and their financial contribution [34]. There are assessment criteria that help to evaluate the quality of the extensions [23, 33]. The evaluation criteria include the customer value offered in the BM, the sustainability of the investment, the stability of the innovation, and the ability to compete with competitors [33]. Moreover, part of the evaluation is whether resources are available or acquisition is possible for the extended BM [24] and whether changes in the relationships and interactions in the service system are needed. The selection will provide the most promising BM extension alternatives and a guide for decision-making [35]. Quick improvements are an important component as well as long-term oriented and promising options [35]. The evaluation is the basis for incremental improvements [23].

The fourth step is associated with the implementation. The process steps include preparatory actions as the actual implementation is not part of the BM design process. Nonetheless, decisions that are made here have an influence on the final design of the BM extensions. Therefore, Figure 1 shows this step by dashed lines. To prepare the implementation, the steps necessary need to be defined. In human-centered design, customer integration should be part of the implementation. Therefore, it should be part of the planning [35]. Results of these activities are documents that specify milestones, structures, finances, and the project plan [23]. Furthermore, the risks of changing the service system associated with the implementation should be documented [24]. In particular, this includes the substitution of humans by ICT [9]. After that, it is possible to follow change management protocols to implement the BM and review the efficacy and any further improvements [37]. The implementation can be monitored by the design parameters and the decision-making and evaluation criteria [34].

6 Application

The developed BM design process was applied in all partner companies. For illustration and better understanding, we demonstrate the use of the process in company C step by step. The partner offers care services that help aged people or persons in other difficult situations to cope with everyday life as independently as possible. Due to the limitation of public and private funding, the care services cannot be applied unlimitedly. When there is little help needed or the need for assistance is at an initial stage, it is not possible to take advantage of these services. To cover this gap and to identify any further demand, the care service company and a housing association started organizing neighborhood assistance in one of their residential properties. To reach more volunteers outside the residential property and to cross-finance the service, the care service company decided to organize the placement of volunteers online.

In the analysis, the existing service has been formally defined for the first time. The results show that the residents are people who usually live alone and have little money from their pension. Some of the people may feel social isolation, loss of importance, and physical decline. They want to experience joy, meet other people, have a structure in their lives, learn, and have the feeling of being needed. The placement service helps older people with early physical handicaps to deal with their everyday life with the assistance of their neighbors. This, including social contacts, helps with simple daily tasks and allows participation in various activities. Neighbors that accompany elderly people can thereby expand their competencies and receive cost compensations due to legal regulations. The ICT-enabled innovation changed the service with the result that the offers and requests can be published and arranged online. Thus, it expanded the group of volunteers to the neighborhood.

In the idea generation of possible BM extensions, options were discussed based on the existing service visualized in a BM canvas. Here, employees of the care service, from the housing association, and representatives of the city authorities discussed ideas and their consequences based on their experiences. Sources for reflections were the needs of the residents and other possible uses of the platform. Obstacles consisted in the clear definition of the users and their needs. Therefore, the prerequisite for adding a new idea was a specified customer group and a corresponding value customers are willing to pay for. In addition, the scalability of the platform was important.

Table 5. Selected Business Model Extensions

Customer	*Value Proposition*	*Revenue Stream*	*Consequences*
Professional service provider	Service platform	Placement fee	Supplementary offering
Local facilities for elderly people	Advertising platform for future customers	Advertising fee	Advertisements for users
City authorities	Execution of municipal tasks	Participation of retirement benefits	Need assessment may be outsourced

In the evaluation and selection, the workshop participants presented the initial idea and summarized the generated BM extensions. Here, a clear definition of the evaluation criteria is crucial for the success. Evaluation criteria were the benchmark from the other volunteering services and the revenue potential. Because competitors already offer similar but specialized services, the BM extensions were evaluated and discussed considering switching costs and a differentiation from competitors. Because the cross-financing of the service was the primary objective, the recurring revenues depicted crucial quality criteria. At the end of the workshop, participants had the opportunity to place their vote for a quick win and a promising long-term solution. For company C, the three BM extensions from Table 5 were selected. These include offers to professional service providers, local facilities for elderly people, and the city authorities. Professional service providers could offer services such as care or medical services for direct booking on the platform. Other local companies that target elderly people could promote their offers on the platform. Also, due to the fact that the platform supports the city authorities in assessing the needs in the neighborhood, the platform provides help for social tasks.

After the workshops, the implementation started with the definition of the steps for each BM extension. Here, this task included the acquisition of the appropriate partners that meet the quality standards, definition of the necessary technical adjustments in the system, and activation of a test group that can provide feedback on the developments. The usage as an advertising platform was a quick win and starting point as the technical adjustments were the smallest and the potential customers easy to reach. Since each placement on the platform also poses a threat to individual privacy, the protection against fraud was a significant risk aspect that needed to be considered for all options. The selected business model extensions enable the care provider to use different revenue streams that can together ensure long-term financing. Commercial success can be achieved even though the main objective has a social and voluntary background.

7 Discussion

Several different approaches aim at generating successful and sustainable HCSSs and corresponding BMs. However, how business models can be adapted to commercialize ICT-enabled service innovations in HCSSs remains unclear. Consolidating contributions from different approaches and considering properties of ICT-enabled HCSSs, we were able to answer the research question of this paper by deriving and testing a BM design process. The objective of the four-step process is enabling those responsible for service innovation in structuring, analyzing, and the decision-making of alternative BM extensions. Although prevailing approaches offer the possibility to design BM extensions, our proposed BM design process provides operationalized steps that allow for the integration of HCSSs, can be used for established service systems, and enable a combination of BM design and service innovations. The value of the method combination is thus included in the in the prescriptive knowledge of the BM design process and its steps.

BM design is a distinctive activity that requires a systematic approach [3]. Nevertheless, present approaches do not comprise the entire process and do not include any operationalization in form of a step-by-step instruction of activities that have to be done. What they offer are, on the one hand, rules that help to find the right design for the right market [24, 33, 35, 38]. These are useful because they make generally accepted statements that are valid for all product and service systems. In addition, they are useful to verify the BM design in terms of strategic fit and competition. They do not have a specific character and do not refer to any specific domain. On the other hand, present approaches offer a high-level process in which BM design represents only a part [8, 34, 37]. These approaches are important in order to consider all of the steps. This does not define a concrete result and does not specify when, what, and how something has to happen in the design process. Other approaches include the right steps but are difficult to use because of their large coverage [23]. Our approach is a guide that includes all steps with specific activities. This is associated with a specific beginning and end of the design process. This allows for an application in different domains and a clear adaptation to the respective needs. In our case, there is a special consideration of HCSSs in the form of an analysis of the interaction and the impact on the people in the service system. The workshops showed that these points were important in the discussion about the suitability and sustainability of the BM design alternatives.

Although BM design is particularly difficult for managers of established service systems [3], no approach provides a guide for this situation. Some approaches are assumed to be able to be applied to already established service systems [23], but they offer no distinction between their actions. The analysis of the environment and competitors [8, 23, 36] provide knowledge that exists in an established service system. These activities focus primarily on gaining competitive advantages, contrary to the focus on commercial success of the ICT-enabled service system [5]. Our proposed BM design process starts with the analysis of the existing service system to promote a shared understanding of the existing service system and its value proposition. In addition, the aspiration of the previous approaches to create new markets [33] is not always practical. Starting with service innovation as a system reconfiguration [18], the ICT-enabled extension of the service system usually aimed at improving existing services. A commercialization on new markets is afflicted with risks [24] and the acquisition of new resources [23], which may not be desirable. Our approach is based on existing technical innovations and/or their ability to create a new value for customers. It turns out that the adaptation of the service system offers the possibility for compliance with restrictions, for example, the focus on interaction, as it is the case in HCSSs.

The development of the process shows that BM design is interdisciplinary as it covers both the design of the service itself as well as related BMs. This is relevant because a service is a process cocreating a context-specific value [13] and HCSSs depend on human interaction and personal services [2]. The value creation and thus the commercialization can only happen with the connection of both. Correspondingly, the literature used in this context comes from various directions. It should be noted that most of the existing approaches focus on the BMs or service. Contributions that include a combination of both are usually of theoretical nature [37]. Our solution provides a BM design that includes the interrelationships between service process and BM innovation, as it has been requested [9]. Different aspects are considered. The process aims at both the customer and the technical innovation. Thus, it covers the most likely drivers of innovation [23]. In addition, BM design provides the possibility to do system-level design and the design of service components [3]. Overall, our approach thus includes different aspects of BM design and service innovation.

The presented research is not without limitations. As BM design is an iterative process [22, 24], the development of the BM extensions with the help of the BM design process is just the first step in achieving a sustainable financing for the ICT-enabled service innovations. We developed and iterated the process with our partners based on a theoretical background. In doing so, new activities were tested in different settings to improve the results. The process presented here includes all activities that have been found useful. Thereby, completeness is not claimed. Nonetheless, we showed that the proposed design process generates several possible and realistic BM extensions that are a good starting point for an iteration with customers and partners. In addition, the acquired knowledge provides an understanding for and from a particular domain with a specific problem. The action research method used to develop the BM design process provides the possibility to learn from joint problem solving, but does not allow an overall generalization of the findings. An abstraction of the findings is only acceptable if the situation and domain are similar to that presented in this paper. To counter this, we based the development of the BM design process on prevailing approaches and iterated the findings in different companies. This allows a comparison with these approaches and expands the application to the presented theoretical background. Therefore, our findings on BM design can be helpful for situations in which service innovation is enabled by ICT and value creation is changed.

8 Conclusion

BM design for ICT-enabled HCSSs focuses on creating value for companies and customers. To capture this value, we derived a BM design process that guides the development of BM extensions. Building on previous contributions from this field, the process was iterated and revised in our action design research project. We used the learnings from three organizational settings to create a systematic and repeatable approach. The result is a BM design process that allows for the revision of the service systems and the corresponding BMs in synergies with and with respect to the existing ICT-enabled HCSSs. The resulting BM extensions foster the commercial success and the long-term financing of the service systems.

With this paper, we offer several contributions. First, the BM design process guides those responsible for service innovation in structuring, analyzing, and the decision-making of alternative BM extensions. The BM design process provides concrete steps and activities and is designed for an application in ICT-enabled HCSSs. Thus, it provides an approach that is relevant to all ICT-related service innovation projects. This process is the first step for the commercial success of ICT-enabled HCSSs. Second, we contribute to the body of knowledge as we provide insights into the combination of the two research streams that have mostly been tackled independently so far, that is, BM design and service innovation. We promote a further understanding of the steps and activities in the design of BM extensions and enable a systematic support of both. Third, by enabling the funding for three volunteer platforms, we solve real-world problems with our research. This confirms the relevance of BM design and shows other objectives besides sustainably competitive advantages.

Future work can build on the developed BM design process. The documentation of the successful application can extend the discussion on BM design and produce suggestions for further improvement. The process can be adapted to other application areas to support the digital transformation. In addition, the long-term success of the developed BM extensions can be examined.

9 Acknowledgements

The research presented in this paper was partially funded by the Active and Assistive Living (AAL) programme and the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number AAL-2013-6-145.

References

1. Barrett, M., Davidson, E., Prabhu, J., Vargo, S.L.: Service innovation in the digital age: Key contributions and future directions. MIS Quarterly. 39, 135–154 (2015).
2. Maglio, P.P., Kwan, S.K., Spohrer, J.: Commentary─Toward a research agenda for human-centered service system innovation. Service Science. 7, 1–10 (2015).
3. Zott, C., Amit, R.: Business model design: An activity system perspective. Long Range Planning. 43, 216–226 (2010).
4. Peters, C., Maglio, P., Badinelli, R., Harmon, R.R., Maull, R.: Emerging digital frontiers for service innovation. Communications of the Association for Information Systems. 39, 136–149 (2016).
5. Storey, C., Cankurtaran, P., Papastathopoulou, P., Hultink, E.J.: Success factors for service innovation: A meta-analysis. Journal of Product Innovation Management. 33, 527–548 (2016).
6. Peters, C., Blohm, I., Leimeister, J.M.: Anatomy of successful business models for complex services: Insights from the telemedicine field. Journal of Management Information Systems. 32, 75–104 (2015).
7. Veit, D., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Kundisch, D., Leimeister, J.M., Loos, P., Spann, M.: Business Models – An Information Systems Research Agenda. Business & Information Systems Engineering. 6, 45–53 (2014).
8. Ebel, P., Bretschneider, U., Leimeister, J.M.: Leveraging virtual business model innovation: A framework for designing business model development tools. Information Systems Journal. 26, 519–550 (2016).
9. Ostrom, A.L., Parasuraman, A., Bowen, D.E., Patricio, L., Voss, C.A.: Service research priorities in a rapidly changing context. Journal of Service Research. 18, 127–159 (2015).
10. Wirtz, B.W., Pistoia, A., Ullrich, S., Göttel, V.: Business models: Origin, development and future research perspectives. Long Range Planning. 49, 36–54 (2016).
11. Kleinschmidt, S., Burkhard, B., Hess, M., Peters, C., Leimeister, J.M.: Towards design principles for aligning human-centered service systems and corresponding business models. In: Proceedings of the 37th International Conference on Information Systems (ICIS). Dublin, Ireland (2016).
12. Osterwalder, A., Pigneur, Y.: Designing business models and similar strategic objects: The contribution of IS. Journal of the Association for Information Systems. 14, 237–244 (2013).
13. Vargo, S.L., Lusch, R.F.: Institutions and axioms: an extension and update of service-dominant logic. Journal of the Academy of Marketing Science. 44, 5–23 (2016).
14. Kleinschmidt, S., Peters, C., Leimeister, J.M.: ICT-enabled service innovation in human-centered service systems: A systematic literature review. In: Proceedings of the 37th International Conference on Information Systems (ICIS). Dublin, Ireland (2016).
15. Rodell, J.B.: Finding meaning through volunteering: Why do employees volunteer and what does it mean for their jobs? Academy of Management Journal. 56, 1274–1294 (2013).

16. Maglio, P.: Editorial—Smart service systems, human-centered service systems, and the mission of service science. Service Science. 7, ii–iii (2015).
17. Breidbach, C.F., Kolb, D.G., Srinivasan, A.: Connectivity in service systems: Does technology-enablement impact the ability of a service system to co-create value? Journal of Service Research. 16, 428–441 (2012).
18. Breidbach, C.F., Maglio, P.P.: A service science perspective on the role of ICT in service innovation. In: ECIS 2015 Research-in-Progress Papers (2015).
19. Maglio, P.P., Spohrer, J.: Fundamentals of service science. Journal of the Academy of Marketing Science. 36, 18–20 (2008).
20. Böhmann, T., Leimeister, J.M., Möslein, K.: Service systems engineering. Business & Information Systems Engineering. 6, 73–79 (2014).
21. Osterwalder, A., Pigneur, Y., Bernarda, G., Smith, A.: Value proposition design: How to create products and services customers want. John Wiley & Sons, Hoboken, NJ (2015).
22. Teece, D.J.: Business models, business strategy and innovation. Long Range Planning. 43, 172–194 (2010).
23. Osterwalder, A., Pigneur, Y.: Business model generation: A handbook for visionaries, game changers, and challengers. John Wiley & Sons, Hoboken, NJ (2010).
24. Chatterjee, S.: Simple rules for designing business models. California Management Review. 55, 97–124 (2013).
25. Blank, S., Dorf, B.: The startup owner's manual: The step-by-step guide for building a great company. K&S Ranch (2012).
26. Ries, E.: The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses. Crown Business (2011).
27. Zomerdijk, L.G., Voss, C.A.: Service design for experience-centric services. Journal of Service Research. 13, 67–82 (2010).
28. Davison, R.M., Martinsons, M.G., Kock, N.: Principles of canonical action research. Information Systems Journal. 14, 65–86 (2004).
29. Baskerville, R.L., Myers, M.D.: Special issue on action research in information systems: Making IS research relevant to practice—Foreword. MIS Quarterly. 28, 329–335 (2004).
30. Baskerville, R.L., Wood-Harper, A.T.: Diversity in information systems action research methods. European Journal of Information Systems. 7, 90–107 (1998).
31. Miles, M.B., Huberman, A.M., Saldaña, J.: Qualitative data analysis: A methods sourcebook. Sage Publications, Thousand Oaks, CA (2014).
32. Kolfschoten, G.L., de Vreede, G.-J.: A design approach for collaboration processes: A multimethod design science study in collaboration engineering. Journal of Management Information Systems. 26, 225–256 (2009).
33. Berry, L.L., Shankar, V., Parish, J.T., Cadwallader, S., Dotzel, T.: Creating new markets through service innovation. MIT Sloan Management Review. 47, 56–63 (2006).
34. Daas, D., Hurkmans, T., Overbeek, S., Bouwman, H.: Developing a decision support system for business model design. Electronic Markets. 23, 251–265 (2013).
35. Kuk, G., Janssen, M.: Assembling infrastructures and business models for service design and innovation. Information Systems Journal. 23, 445–469 (2013).
36. Solaimani, S., Guldemond, N., Bouwman, H.: Dynamic stakeholder interaction analysis: Innovative smart living design cases. Electronic Markets. 23, 317–328 (2013).
37. Chew, E.K.: iSIM: An integrated design method for commercializing service innovation. Information Systems Frontiers. 18, 457–478 (2016).
38. Giessmann, A., Legner, C.: Designing business models for cloud platforms. Information Systems Journal. 26, 551–579 (2016).

The Impact of Blockchain Technology on Business Models in the Payments Industry

Friedrich Holotiuk[1], Francesco Pisani[1], and Jürgen Moormann[1]

[1] Frankfurt School of Finance & Management, ProcessLab, Frankfurt a. M., Germany
{f.holotiuk,f.pisani,j.moormann}@fs.de

Abstract. Because of its potentially disruptive influence on business models (BMs), blockchain technology has sparked a lively debate among researchers. Our Delphi study sets out to explore the impact of blockchain in payments, which represents a major cornerstone of banking and the cradle of this technology. The results, grouped around four areas of thoughts, indicate that blockchain allows the offering of new services and renders some of the current ones obsolete. This consequently impacts the financial structure of firms in the payments industry and further generates great potential for new BMs while making some existing ones obsolete. Eventually, new players, which are better able to leverage the potential of blockchain, will give a strong impulse to this development. Our findings contribute to the literature by providing new insights about the impact of innovative technologies on BMs and have further practical implications by presenting a better understanding of future BMs in payments.

Keywords: *blockchain, business model, Delphi method, innovation, payments industry*

1 Introduction

Technological changes pose new challenges and generate further opportunities for firms. In particular, innovative technologies have the potential to modify the equilibrium among the firms in an industry. Leading firms consistently fail to stay at the top of their industry when technological discontinuities occur [1]. Not promptly identifying their impact on the business model (BM) may even result in a ruinous error [2]. Examples are the introduction of digital cameras, smartphones, and online streaming. Companies such as Eastman Kodak, Nokia, and Blockbuster had to leave the market because of their inability to adapt their BMs to a changed technological environment. Hence, it is extremely important to clearly assess the consequences that the introduction of new technologies for the BMs in the affected industry can have.

During the last years, the financial services sector has gone through far-reaching changes partly due to the recent financial market crisis. Nowadays, the move toward digitalization of processes and products is further pushing banks and other financial institutions to rethink their strategies, BMs, and operations. The advent of new technologies, combined with a decline in margins and the rise of new competitors, are pressuring incumbent companies to find viable solutions that would allow them to cope with

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Holotiuk, F.; Pisani, F.; Moormann, J. (2017): The Impact of Blockchain Technology on Business Models in the Payments Industry, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 912-926

the new environment. It is under this scenario that especially a technology named blockchain is attracting the attention of the actors in the financial sector for its potentially revolutionary enhancements of operations and financials. Following [3], blockchain technology stands for a radical shift to direct transactions between end parties without current intermediary services building on a consensus mechanism to verify new transactions, and a decentralized record keeping of all transactions.

Although the misuse of bitcoins has aroused some skepticism in the payments industry, the technology behind this early form of new digital currency has gradually imposed its evocative presence. Over recent years, blockchain-based applications have multiplied and use cases that cross over the boundaries of the payments field have been envisioned. Blockchain is thought to have an extraordinary potential [4,5] and its adoption in the payments industry is believed to be groundbreaking [6].

Nevertheless, even if preliminary predictions about the effects of blockchain for specific parts of the financial sector have been made, a clear delineation of the consequences that it might have for the overall payments industry has yet to be identified. To address this point and contribute to the literature related to BMs and technological innovations, this article answers the following research questions: *How does blockchain technology impact current and new business models in the payments industry?*

The results are based on a Delphi study among experts in the payments industry knowledgeable about blockchain technology. The study is composed of three rounds. In total, 45 experts from several European countries have agreed to take part in the study. The results, grouped around four areas of thoughts, indicate that blockchain technology will affect the BMs in the payments industry by (1) allowing new services and making some of the current services obsolete. (2) Through this change in services, a subsequent impact on the financial structure of firms in the payments industry is realized. (3) This generates a great potential for new BMs in the market while some existing ones become obsolete. (4) Eventually, the industry is impacted by new players that are better able to leverage the potential of the technology. Our contribution lies in the analysis of the impact of new technologies on BMs using the example of blockchain technology.

The remaining part of this article is organized as follows. Section 2 outlines the conceptual background of the paper. In particular, it covers cornerstones of BMs as well as blockchain technology and delineates the current situation of the payments industry. Section 3 explains the methodology of our study, whereas section 4 presents the results of the Delphi analysis. Section 5 discusses the implications and limitations of the study. Section 6 concludes the paper.

2 Background

2.1 Business Models

BMs are a relatively new concept in management studies [7]. The first apparition of the term can be dated back to 1957, but it is not until the end of the twentieth century that it has broadly attracted the attention of researchers [8]. Although a specific definition has still to be found [9], a BM has been identified as the "story that explains how an enterprise works" [10] or also as the way firms do business. Further definitions of the term have been proposed in the past. Nevertheless, the concept of value represented a central aspect in many of the descriptions given by researchers since the early studies published around the topic [9–13]. Therefore, the definition of BM adopted in this paper follows this approach established around the concept of value by Osterwalder and Pigneur [12], which identifies a BM as "the rationale of how an organization creates, delivers, and captures value".

Next to the concepts of value, strategies, and processes, research on BMs has also delineated a strong link with technology. In particular, the design of IT systems able to support BMs has been one the most cited issues in the early 2000s [14]. BMs represent an intermediate layer – the link between a firm's strategy [8], processes [15], and information technology (IT) [7,16]. This relationship between IT and BMs is fundamental to understand, design and leverage organizations [16]. Technological innovations alone do not assure the success of a firm [17]. A BM is important to secure a competitive advantage [18] and to mediate between the development of technologies and the creation of economic value [13] as well as firm performance [19]. It is essential to capture the value from an innovation and assure its commercial success [11].

The link between BMs and technology further assumes a particular relevance when analyzing the introduction of technological innovations in a specific industry. Inertly maintaining longstanding beliefs and not adapting a company's BM to technological discontinuities has proven to carry fatal consequences [2]. Evidence from the drug industry suggests that in order to have such an effect, new technologies should "create new dependency ties and reshape collaboration patterns" [20]. Therefore, it is imperative to assess opportunities and challenges engendering from technological changes. This article uses the blockchain in payments to evaluate the impact that a new technology has on the BMs of firms in this specific field.

2.2 Blockchain Technology

Blockchain has been initially launched as an approach to payment transactions based on cryptography to provide an alternative mechanism for the trust between two transacting parties [21]. This technology enables a collective bookkeeping system (ledger), which, by means of a mathematical function (hash function), allows participants to reach an agreement on the approval of a transaction. The information concerning single transactions is gathered in 'blocks'. These blocks are reviewed and verified by the network and added in a chronological order on the computers of all participants of the network. A distributed ledger of verified transactions of a particular unit is then provided to the network. As such, the traditional role played by financial institutions as trusted third party, able to mitigate the risk behind a transaction, is under scrutiny.

Bitcoin was the first digital currency and remains the largest until now [22]. Furthermore, it represents one of the most famous applications of blockchain technology. Nowadays, blockchain is being proposed as a solution for a broad spectrum of transactions, which range from real-time payments between two parties (rapid settlement and without requiring a bank account) to transferring funds across currencies (micro payments, remittances), and digital assets (digitally stored records of ownership).

The impact of blockchain technology might go much further than some modified processes and a few new products and services. A number of authors expect that the consequences might be much further reaching in that entire BMs will be affected [4,5]. In this sense, the impact on BMs through blockchain technology might be a good example for the far-reaching potential of IT [23].

Accordingly, blockchain technology or the more general term of distributed ledger technology (DLT) has raised huge interest in the Information Systems (IS) community, e.g. with regard to trust and cryptographic aspects [24], to the procedure and implications [3], as well as to diverse issues of virtual currencies [25]. Hence, this paper aims to further contribute to this literature by providing an enhanced understanding of the implications that the introduction of this technology might have on BMs.

2.3 Payments Industry

The payments industry represents one of the major business fields of financial institutions. In effect, payments are not only a lucrative source of revenues, but they are the anchor product for various other services and, furthermore, a critical element in terms of customer data. For banks, payment information is a source of knowledge about the customer and, further, is an opportunity to generate points of reference to integrate business processes into the processes of their customers. Thus, losing stakes in payment transactions would cause disastrous consequences for banks.

Currently, the payments industry in Europe finds itself in a state of great upheaval triggered by regulatory as well as political initiatives. Among them it is worth mentioning the creation of the Single Euro Payments Area (SEPA), the establishment of instant payments, which is already decided and on the way, the revised payment services directive (PSD2), which will become effective in 2017 in all EU member states, and the regulation on interchange fees (EU 2015/751). Most of the current projects serve the goal to harmonize the euro payments market in Europe, as well as to encourage more competition and open the market to new entrants.

Globally, the emergence of smartphones has allowed new players, such as large technology and telecommunication enterprises, entering the market. Furthermore, numerous companies from the fintechs arena (start-up companies in the financial services sector relying heavily on IT) have emerged.

Squeezed in between the need for investments in compliance and IT, the erosion of income from traditional sources, and increased competition, the BM of many financial institutions is already under pressure. Therefore, any further attempt to make the current payment infrastructure obsolete or to pull away payment transactions from banks and other financial institutions strongly contributes to deteriorating their BMs. In this regard, blockchain technology represents a fearful threat, especially since it might switch off the third-party function of financial institutions in payments. At the same time, however, the reduction of costs that could be realized by the use of blockchain in payments induces financial institutions to closely look at its development.

This promising potential of blockchain technology has roused large attention at existing payment infrastructure operators such as SWIFT, providers of international payment transactions as well as regulators. Enterprises from both technology as well as the financial services sector are considering and launching prototypes of blockchain-based solutions. In particular, incumbent companies try to defend their BMs by applying a range of strategies from developing in-house platforms to directly investing in blockchain companies, partnering with them, or offering accelerator services to explore blockchain applications.

Large banks have started to participate in worldwide collaborations (e.g. R3, among them Citibank, Credit Suisse, Deutsche Bank), and almost all major consultancies as well as auditing firms offer their expertise and try to position themselves as the leading knowledge carriers.

3 Research Method

The analysis at the center of this paper is based on a Delphi study conducted among experts from the payments industry knowledgeable of the blockchain technology. Given the lack of existing research and the exploratory nature of our study, open qualitative interviews would have been an option. However, the industry still shows a high degree of uncertainty on the study's topic. Furthermore, based on our industry insight, specific expertise could clearly be located. This advised a multi-stage study in a more formalized and group-oriented approach. Therefore, the Delphi approach was the method of choice [26]. The Delphi method was developed in the 1950s [27] and has become a common tool for measuring and aiding forecasting and decision-making [26]. It is especially appropriate for exploratory theory building on interdisciplinary issues, which often involves new or future trends [28,29]. Hence, the method is highly recognized in research concerning technology forecasting [30,31] and has been used extensively in IS research to identify and rank key issues for management action [32].

The Delphi method allows for the discussion of a complex issue through a structured communication process [33]. Dakey and Helmer [27] define Delphi as a method that attempts to obtain the most reliable consensus of a group of anonymous experts. Four distinct characteristics are presented by von der Gracht [34]: anonymity, iteration, controlled feedback, and statistical group response. With respect to our research aim and as suggested by Murry and Hammons [35] we chose a 3-round procedure. In this regard, we follow Fan and Cheng [36], who suggest three rounds as being sufficient to reach consensus and borne in mind time constraints which might influence the method [34].

Round one (R1) aimed to derive panelists' insights and opinions. In round two (R2) panelists evaluated the results of R1. Round three (R3) asked panelists to reevaluate the results in light of the group feedback. Although we recognize that the Delphi method has been widely reviewed [31,33,37], we briefly outline the identification of experts, data collection as well as analysis, and explain the specifics of our study.

3.1 Identifying the Panel of Experts

The most important criterion when selecting panelists is the individual expertise on the issue under study [38]. Therefore, we took the requirements described by Hill and Fowles [39] as well as Adler and Ziglio [31] into account. Accordingly, we selected qualified experts depending on their work experience in payments and/or blockchain technology, their professional position, and the role and background of the company they work with. A key requirement for experts to be selected was a thorough understanding of blockchain technology to assess its implications on payments. In addition, a deep understanding of payments was needed to assess industry-specific consequences for BMs. For the identification and validation of experts, we used web search, talks with practitioners, and databases of professional networks. Hence, the Delphi panel was composed to be a representative mix of experts [3] and included 45 panelists: 16 (35%) from consulting, 11 (24%) from fintechs, 6 (13%) from banks, 4 (9%) from academia, 3 (7%) from public institutions, 3 (7%) from payment service providers, and 2 (4%) technology providers. The high number of consultants is explained by their current leading role in collaborative projects with banks and technology companies with regard to blockchain technology. The panel has not changed throughout the study, but size reduced due to minor dropouts. The stable core enabled us to deduce a broad range of answers from a wide spectrum of organizations while still staying with a clearly focused evaluation and consensus process [40].

3.2 Data Collection and Analysis

Due to the iterative and multi-stage nature of Delphi studies, data collection and analysis are presented jointly. In R1, we sent out 45 emails to the panelists where we asked them to independently provide ideas, thoughts, and opinions on the development of blockchain technology. According to Linstone and Turoff [33], we designed R1 with an open-end format, suggesting starting points around BMs. This was done to elicit individual perspectives, judgments, and opinions from each panelist. In order to develop a general framework in the direction of our research question, the starting points were created by the researchers as suggested by Schmidt [32]. Hence, to stimulate answers in R1, broad questions (e.g. future scenarios, products, and technology) were provided where panelist could deliver their input. All answers were submitted via an online form.

We received 38 responses. For easier reading and analysis the responses were collected in one document resulting in 20,000 words of qualitative data. In order to distil the most relevant statements, the input was coded by three independent researchers with a moderator coordinating the coding activities. First, the researchers went through all answers and developed their own code list. Second, the moderator guided the discussion among the researchers to generate one code list which reflected all relevant input. Finally, the researchers translated all codes into better readable and easily understandable statements. As an example, the code “new business models” was translated to “With the blockchain technology new business models in payments will develop”. All statements relevant to our research are presented in section 4. By means of the coding in R1, an initial set of 45 statements was produced describing the implications of blockchain technology in payments.

In this paper, we analyze and discuss those statements which are relevant for BM research in relation to blockchain technology. The researchers identified 17 out of the 45 statements as being relevant to the objectives of this research. The statement selection was based on the following criteria: threats and opportunities for existing BMs, need for revising current BMs, implications for designing new BMs, and new service offerings in the industry with substantial potential for new BMs.

For the subsequent evaluation of the statements in R2 and R3, we had to take into account that the expert panel consists of practitioners with limited time as well as relatively low methodical understanding. Hence, to better facilitate the evaluation, the statements were presented through the use of an online tool (Qualtrics) with a strong focus on intuitive readability. In R2, we exclusively considered the 38 panelists who completed R1. These experts were presented with the statements generated in R1 and asked to provide an evaluation of each statement on a six-point Likert scale ranging from "Strongly agree" to "Strongly disagree". Six points were chosen to encourage clear decisions toward agreement or disagreement but at the same time to offer enough options for a differentiated evaluation. At the end of R2, the evaluation of each statement was received from 36 out of the 38 panelists.

This group of 36 experts was further considered in R3, where the identical statements from R2 were presented to the panelists, along with the group's responses from R2 combined with each panelist's own evaluation. Since we required a high degree of clarity to present the responses we adopted intuitively usable measurements. Hence, solely graphical representations of the evaluations were shown. This approach exceeds the standard Delphi method, but assures the correct interpretation, as Argyrous [41] stresses that the mean of ordinal data is misleading and incorrect. In the end, panelists were asked to provide their individual evaluations in light of the group evaluations in R2. In total, 34 responses were collected from R3 (Table 1).

Table 1. Response rates within the Delphi panel

Round 1		*Round 2*		*Round 3*	
Sent out	Complete responses	Sent out	Complete responses	Sent out	Complete responses
45	38 (84.4%)	38	36 (94.7%)	36	34 (94.4%)

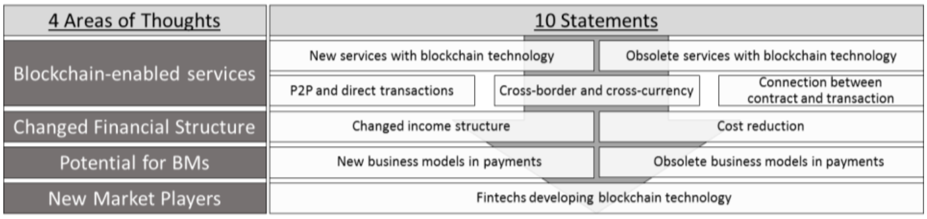
After finishing R3 we checked group stability, as defined by Dajani et al. [42] and Linstone and Turoff [33], with the majority of panelists agreeing to the statements. Across all statements, the average for agreement was 87% and only 13% for disagreement.

Next, we compared two statistical measures, variance and variation, of R2 and R3 to determine if consensus was achieved. The average variance was reduced from 1.23 in R2 to 0.96 in R3. Furthermore, the average variation decreased from 47% in R2 to 43% in R3. Finally, we selected those statements of the initial 17 with the highest consensus values. First, we used a predefined level of agreement of 75% on our 6-point Likert scale. This seems reasonable as similar research uses percentages between 60% [43] and 80% on a 5-point Likert scale [44]. Second, we required a variation score below 50% as suggested by English and Keran [45]. Third, statements were excluded when the variance was above 1.0 [34]. As a result, we were able to identify ten statements meeting the before mentioned criteria.

4 Results

The ten statements are the result of the Delphi method and best summarize the implications of blockchain technology on BMs in the payments industry based on the expert panel. Figure 1 illustrates how the ten statements are synthesized into four areas of thoughts: (1) Blockchain-enabled services as a first cluster indicate how new services around peer-to-peer (P2P) and direct transactions, cross-border and cross-currency transactions, as well as the connection between contracts and transactions are being introduced. At the same time some existing services are rendered obsolete. (2) This change in services causes a change in the financial structure of firms in the payments industry. (3) As a consequence, there is a great potential for new BMs in the market while some existing ones become obsolete. (4) A strong impulse to new BMs is given by new players like fintechs, which are better able to leverage the potential of blockchain technology. Details on the opinions of the panelists are provided in the following.

Figure 1. Implications of Blockchain Technology for BMs in the Payments Industry



We see a strong consensus around the impact on payment services due to the introduction of blockchain, and we argue that there are direct implications at the BM level as the design of BMs involves the definition of services a firm delivers [11].

On the one hand, panelists stress that the development of blockchain technology allows new service offerings to be brought to the market. In more detail, experts mention three service areas, which play a major role in the further development. These services are shaping the development of BMs and are forerunners of the change to come in payments: (1) Blockchain technology is expected to make direct transactions possible without any third party acting as "*trust agent*"[1]. Hence, "*transaction can be executed peer-to-peer*" directly between two contractual parties (peers). P2P transactions can occur between identified parties such as firms or customers; but also between unidentified parties like machines (cars etc.) or even unbanked customers. Furthermore, "*transactions without a middle man*" are paving the way for decentralized trading markets. (2) Blockchain technology is thought to improve international transactions in

[1] All direct citations in this section are taken from the answers panelists provided in R1 and are formatted italic.

cross-currency and cross-border context. The huge potential of these improvements become obvious when looking at globalized trade and the high inefficiency of the current global payment infrastructure. Today, cross-border transactions are time-consuming, lengthy, and expensive. Blockchain technology will make these payments "*faster and cheaper*", i.e. faster by providing a solid, common infrastructure across borders for transactions, and cheaper by removing expensive intermediaries, thus overcoming today's "*lack of trust*". If blockchain technology allows easy international transactions based on digital currencies, currency exchange will erode as a service and remove pricey currency exchange offices. Furthermore, due to the inclusive nature of the technology, global and permissionless accessibility, current high charges for remittances by third parties will fade and erode the respective BMs as individuals can participate directly in remittances abroad. These improvements will be some of the "*biggest impacts*" of blockchain technology. (3) A completely new service blockchain technology will allow is the connection between contracts and transactions. Hence, the technology can be used to keep records of "*contracts of purchase and passing of property*" in addition to the actual transaction. Thereby, contracts of purchase can be directly linked to payment transactions, which is referred to as smart contracts. As a result, blockchain technology can be used as a "*proof of ownership*" as well as a proof of payment. The development of smart contracts will allow the "*automated execution of transactions*". Hence, smart contracts prove to be a critical cornerstone in the current advancements around the internet of things. Finally, the connection between contracts and transactions allows 'programmable' money flows and automation of transactions, which leads to decentralized autonomous organizations, where business rules are coded in the organization and executed automatically under certain conditions [46]. Further, extended service offerings mentioned in several answers, touch upon the relevance of "*making money out of data intelligence*" and data in general. The future for market players will be around payment services enhancing the traditional transaction services. Data can be used to offer "*data analytics*" to deliver deeper insights into payments, which contributes to enhanced "*fraud detection and prevention*". Other important services, which will be needed, are the conversion between traditional payments and blockchain payments as well as personal financial management.

On the other hand, blockchain technology is expected to render obsolete current payment services like third-party trust service, clearing and settlement, as well as reconciliation. As a starting point, most panelists mentioned that today's processes are "*inefficient and slow"*. They particularly refer to the current payment infrastructure (SWIFT and SEPA transactions), which require a lot of manual steps and, hence, *"transfers at a relatively high cost"*. Due to the unified record keeping in the blocks, clearing and settlement services will no longer be needed for payments based on blockchain which rather leads to the implementation of "*fully automated reconciliation*". As a result, the omission of entire process steps is expected to eliminate core services of existing BMs, questioning their existence. Ultimately, blockchain technology *"will allow equal access"* for market players, making payments a commodity.

As a consequence of new services, the financial structure of the BMs changes. The implications are twofold: On the one hand, the revenue structure resulting from payment transactions changes substantially, which means that traditional sources of revenue die out. Yet, at the same time new ones emerge. There is consensus that "*payments will be a commodity*" resulting in very low margins. Furthermore, the currently mainly margin-based revenue structure will erode with transaction fees dropping to "*even less than cents*". Current margins benefit from high complexity and artificially created boundaries between payment networks, which will vanish with blockchain technology. Revenue streams of BMs have to be shifted away from transaction-based margins and have to focus on the provision of "*user-friendly and secure platforms*" or the management of smart contracts. On the other hand, blockchain technology allows cost reductions. For example, the replacement of the currently inefficient payment infrastructure by blockchain technology will free up capital. Also the costs for processing transactions drop, making the transfer of money cheaper. "*The opening of formerly closed systems*" provides great potential to reduce costs. Overall, the increase of efficiency will "*address the rising costs*" of regulation and allow more efficient compliance due to increased transparency. For instance, "*know-your-customer processes will be streamlined*", which results in decreased costs. Furthermore, the faster execution of transactions leads to a reduced risk of default and, hence, to lower costs.

Building on the blockchain-enabled services and the changed financial structure, strong effects on BMs are observable and play a major role when discussing blockchain technology. Nonetheless, the perception is double-edged. On the one hand, we see a strong consensus that new BMs with regard to payments will emerge. For example, panelists stress the importance of data by underlying "*data analytics and further data-related services*". This is in accordance with the trend in our research, that payments-related BMs will only survive if new services are added like "*payments-extending services and products*" and thus BMs are enhanced. Only the creation of "*value-added service*", complementing current BMs, will allow financial institutions to keep their customer base stable. For example, panelists point out that future BMs will no longer build on account service fees but "*hosting and data security fees*" and will be able to "*monetize interfaces*", not just services.

Quite contrary to the great potential of blockchain technology for payments, we see, on the other hand, an equally strong consensus that some BMs in payments will become obsolete. Examples are the traditional margin-based, intermediary or trusted party BMs. The role of a trustworthy broker ("man in the middle*") "will be redundant with blockchains*". Intermediaries face the problem of complete eradication as they are going to be "*extinct because their BM is being replaced with a more efficient mechanism*". Margins cannot provide a source of revenue, as the mere execution of transactions will lose importance in blockchain systems. Blockchain features, like direct transactions, speak against the current structure of the market. This phenomenon is also recently discussed in academic research [3]. Furthermore, it is questionable if financial institutions can maintain their current function as trusted party as advancements with blockchain will enable features like direct transactions and equal access to the market for all participants. It is noteworthy that participants compare the "*future role of payment service providers to the letter mail in the age of the internet*".

Finally, the described changes and implications give rise to new market players. As new players, and particularly fintechs, enter the market, new BMs are expected. Following the panelists, fintechs will play an important role in the context of blockchain technology application in payments. Panelists see fintechs as an "*enabler for market infrastructure*" and as "*specialized providers from outside with a catalytic role*". The increasing number of fintechs like Ethereum and Ripple supports this view. Moreover, there are certain structural and technological boundaries in existing financial institutions that make it hard to change the underlying technology the business is running on (e.g. back office software, inter-organizational payment networks, supra-authority infrastructure). In contrast, fintechs have the advantage of being able to decide for a new technology with fewer dependencies and, hence, adopt blockchain technology considerably quicker. They will play out their advantage to occupy parts of the value chain and offer services industry-wide, which will force existing players to "*acquire white label blockchain solutions*" from fintechs to stay in the market.

5 Discussion

This paper is motivated by the debate on the influence of blockchain technology [24,47] and the growing body of literature on cryptocurrencies [25], combined with the necessity to assess how this would impact BMs in the payments industry.

Our findings indicate that changes due to the introduction of blockchain are reflected in new services as well as new revenue structures and eventually new BMs. Following the definition by Christensen et al. [1], blockchain technology exemplifies disruptive market capabilities, as it currently offers features, which seem uncommon or less convenient, but will change and impact the industry in the long run. The disruptive power of blockchain is further supported by previous literature on the topic [47]. Our findings add a new perspective to the literature on BMs by showing how a new technology could actually impact the BMs of firms in those industries where it is introduced. In this regard, we extend the insights provided by Sabatier et al. [20] and suggest that new technologies have the potential to disrupt the equilibria within an industry especially by undermining the service logic and the revenue structures established within this industry. To better address these situations, banks and other financial institutions in the payments industry have to rethink their current BMs and allow for experimentation. Based on the ongoing development of the technology, these firms have to better assess implications on their current services and products to prepare for the arrival of blockchains. Through cooperation with fintechs, incumbent financial institutions could be better able to benefit from the fintechs' dynamism while limiting the need for large and risky investments until a clear path for the development of the technology has emerged. As first cross-currency transaction conducted by some international banks and the fintech Ripple indicates, a couple of financial institutions have already adopted such an approach.

With the elimination of process steps, the impact of technology will also be reflected in the firms' business processes. Furthermore, we acknowledge the transformative power the technology also poses on society. The inclusive character of blockchain, equal access for all participants and almost zero participants costs, allows addressing the unbanked customer in the developed as well as developing part of the world and new BMs can offer a wide range of services to those customers [48].

The technology is still developing and a lot of try-outs, prototypes and experiments are needed. Nonetheless, the technology receives a lot of attention and was recently added to Gartner's hype cycle of emerging technologies in 2016. Interestingly, blockchain is placed at the peak of the cycle, which shows the attention it has raised as well as the inflated expectations it faces. Still, moving along the hype cycle does not mean that the technology is a temporary hype as some critics raised in the academic literature might suggest [49]. Instead, the development shows the different stages technologies are going through and our findings do not refer to temporary effects. Nonetheless, the full potential of the technology is still not completely foreseeable and the application of blockchain is still highly context-specific. Not every setting in the payments industry is suitable and a number of questions remain unanswered [3]. Questions raised by our panel of experts such as "*For how long shall systems work in parallel?*" or "*Will the technology prove successful and actually make things better?*" still remain open.

Limitations of our research include the reporting of primarily positive aspects of blockchain technology by the experts in R1. Hence, positive aspects seem to outweigh, which can be explained by the fact that people naturally tend to report aspects they are aware of or agree with. Thus the statements focus less on possible drawbacks. Experts were asked for their opinions and judgments on blockchain technology, which might sometimes be far from real use cases or first prototypes. Not every expert necessarily has gathered personal experience with the technology. Furthermore, the expertise on this new technology is still immature and uncertainty remains.

Blockchain technology is at a young age and research on the matter is still scarce. Future research should deepen the findings of this paper in two directions. First, as described by Al-Debei et al. [15], we see implications from BMs to the underlying business processes. Hence, an analysis of business processes at an intra-firm level is promising and allows studying implementations of blockchain technology within fintechs or first prototypes developed by incumbents. This could also be deepened with a case study analysis and/or interviews with founders of fintechs with the focus on blockchain. Second, there are interactions between the BM and the overlying business strategy [15]. Therefore blockchain technology, as it impacts BMs, also yields strategic implications. Strategy has to include digital technology and to establish a closer link between business and IT [50]. Therefore, studying the impact of blockchain technology on BMs and corporate strategy will provide a better understanding of the fusion of business and IT strategy [50].

Our results led to a research agenda in the field of blockchain technology. First of all, it appears to be decisive to better understand new, customer-centered services enabled by blockchain technology and how these services could be used in existing and new BMs in the field of payments. An analysis of the services will also allow to investigate the interplay between new and existing players. Next, it is fundamental to analyze the changes in the cost structure and, hence, the financial benefits of blockchain as costs represent a major driver for the adoption of new technology and changes to the BM. From the adoption and integration of blockchain technology, researchers can deduce the adaptions needed for existing BMs and the potential for new BMs.

6 Conclusion

Blockchain is a new technology with potentially disruptive power, which yields implications for a number of industries [47]. First applications arose in the financial services sector with bitcoins [21], which puts the payments industry at the center of innovations around blockchain technology. Our study is the first of its kind to gather a high number of experts and gain a better understanding of the implications on BMs in the payments industry. Our paper outlines the changes due to blockchain technology, which are clustered in four areas of thoughts. First, new services are introduced, which foster P2P transactions, cross-border and cross-currency transactions, as well as the connection between contracts and transactions, and, hence, make current services obsolete. As a consequence, financial structures of existing BM will change. Third, these changes will be reflected in the development of new BMs, making some existing BMs obsolete. Finally, these changes create a potential for fintechs to enter the market by leveraging blockchain technology.

Summarized, our research delivers insights into how changes in payments, due to blockchain technology, progress and in what directions firms have to think to overhaul their BMs. Our research contributes to BM literature by analyzing the impact of new technologies. Furthermore, the findings yield new research avenues, which are promising to further explore the topic of blockchain. In the end, the saying "one secret to maintaining a thriving business is recognizing when it needs a fundamental change" [18] might prove right once again with BMs in payments.

References

1. Christensen, C.M., Raynor, M. & McDonald, R., 2015. What is Disruptive Innovation? *Harvard Business Review*, 93(10), pp.44–53.
2. Tripsas, M. & Gavetti, G., 2000. Capabilities, Cognition, and Inertia: Evidence From Digital Imaging. *Strategic Management Journal*, 21(10), pp.1147–1161.
3. Roßbach, P., 2016. Blockchain-Technologien und ihre Implikationen. *BIT – Banking and Information Technology*, 56(1), pp.54–69.
4. Tapscott, D. & Tapscott, A., 2016. *Blockchain Revolution: How the Technology Behind Bitcoin Is Changing Money, Business and the World*, New York: Penguin.
5. Swan, M., 2015. *Blockchain: Blueprint for a New Economy*, Sebastopol: O'Reilly.
6. Bott, J. & Milkau, U., 2016. Towards a Framework for the Evaluation and Design of Distributed Ledger Technologies in Banking and Payments. *Journal of Payments Strategy & Systems*, 10(2), pp.153–171.
7. Al-Debei, M.M. & Avison, D., 2010. Developing a Unified Framework of the Business Model Concept. *European Journal of Information Systems*, 19(3), pp.359–376.
8. Osterwalder, A., Pigneur, Y. & Tucci, C.L., 2005. Clarifying Business Models: Origins, Present, and Future of the Concept. *Communications of the Association for Information Systems*, 15(1), pp.1–43.
9. Wirtz, B.W., Pistoia, A., Ullrich, S. & Göttel, V., 2016. Business Models: Origin, Development and Future Research Perspectives. *Long Range Planning*, 49(1), pp.36–54.
10. Magretta, J., 2002. Why Business Models Matter. *Harvard Business Review*, 80(5), pp.86–87.

11. Teece, D.J., 2010. Business Models, Business Strategy and Innovation. *Long Range Planning*, 43(2–3), pp.172–194.
12. Osterwalder, A. & Pigneur, Y., 2010. *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*, New York: John Wiley & Sons.
13. Chesbrough, H. & Rosenbloom, R.S., 2002. The Role of the Business Model in Capturing Value From Innovation: Evidence From Xerox Corporation's Technology Spin-off Companies. *Industrial and Corporate Change*, 11(3), pp.529–555.
14. Pateli, A.G. & Giaglis, G.M., 2004. A Research Framework for Analyzing eBusiness Models. *European Journal of Information Systems*, 13(4), pp.302–314.
15. Al-Debei, M.M., El-Haddadeh, R. & Avison, D., 2008. Defining the Business Model in the New World of Digital Business. In *AMCIS 2008 Proceedings*. pp. 1–11.
16. Veit, D., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Kundisch, D., Leimeister, J., Loos, P. & Spann, M., 2014. Business Models. *Business & Information Systems Engineering*, 6(1), pp.45–53.
17. Zott, C., Amit, R. & Massa, L., 2011. The Business Model: Recent Developments and Future Research. *Journal of Management*, 37(4), pp.1019–1042.
18. Johnson, M.W., Christensen, C.M. & Kagermann, H., 2008. Reinventing Your Business Model. *Harvard Business Review*, 86(10), pp.57–68.
19. Baden-Fuller, C. & Haefliger, S., 2013. Business Models and Technological Innovation. *Long Range Planning*, 46(6), pp.419–426.
20. Sabatier, V., Craig-Kennard, A. & Mangematin, V., 2012. When Technological Discontinuities and Disruptive Business Models Challenge Dominant Industry Logics: Insights from the Drugs Industry. *Technological Forecasting and Social Change*, 79(5), pp.949–962.
21. Nakamoto, S., 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. Available at: https://bitcoin.org/bitcoin.pdf [Accessed July 25, 2016].
22. Robleh, A., Barrdear, J., Clews, R. & Southgate, J., 2014. Innovations in Payment Technologies and the Emergence of Digital Currencies. *Bank of England*. Available at: http://www.bankofengland.co.uk/publications/Documents/quarterlybulletin/2014/qb14q3digitalcurrenciesbitcoin1.pdf [Accessed July 10, 2016].
23. Brynjolfsson, E. & McAfee, A., 2014. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, New York: W. W. Norton & Company.
24. Beck, R., Stenum Czespluch, J., Lollike, N. & Malone, S., 2016. Blockchain – the Gateway To Trust-Free Cryptographic Transactions. In *ECIS 2016 Proceedings*.
25. Kazan, E., Tan, C.-W. & Lim, E.T.K., 2015. Value Creation in Cryptocurrency Networks: Towards A Taxonomy of Digital Business Models for Bitcoin Companies. In *PACIS 2015 Proceedings*.
26. Rowe, G. & Wright, G., 1999. The Delphi Technique as a Forecasting Tool: Issues and Analysis. *International Journal of Forecasting*, 15(4), pp.353–375.
27. Dalkey, N. & Helmer, O., 1963. An Experimental Application of the Delphi Method to the Use of Experts Author. *Management Science*, 9(3), pp.458–467.
28. Meredith, J.R., Raturi, A., Amoako-Gyampah, K. & Kaplan, B., 1989. Alternative Research Paradigms in Operations. *Journal of Operations Management*, 8(4), pp.297–326.
29. Akkermans, H.A., Bogerd, P., Yücesan, E. & van Wassenhove, L.N., 2003. The Impact of ERP on Supply Chain Management: Exploratory Findings from a European Delphi Study. *European Journal of Operational Research*, 146(2), pp.284–301.
30. Turoff, M., 1971. Delphi and its Potential Impact on Information Systems. In *AFIPS 1971 Proceedings*. pp. 317–326.

31. Adler, M. & Ziglio, E., 1996. *Gazing into the Oracle: The Delphi Method and its Application to Social Policy and Public Health*, London: Jessica Kingsley.
32. Schmidt, R.C., 1997. Managing Delphi Surveys Using Nonparametric Statistical Techniques. *Decision Sciences*, 28(3), pp.763–774.
33. Linstone, H.A. & Turoff, M., 1975. *The Delphi Method: Techniques and Applications*, Reading: Addison-Wesley.
34. von der Gracht, H.A., 2012. Consensus Measurement in Delphi Studies: Review and Implications for Future Quality Assurance. *Technological Forecasting and Social Change*, 79(8), pp.1525–1536.
35. Murry, J. & James, H., 1995. Delphi: A Versatile Methodology for Conducting Qualitative Research. *The Review of Higher Education*, 18(4), pp.423–436.
36. Fan, C.K. & Cheng, C.-L., 2006. A Study to Identify the Training Needs of Life Insurance Sales Representatives in Taiwan Using the Delphi Approach. *International Journal of Training and Development*, 10(3), pp.212–226.
37. Delbecq, A.L., Van de Ven, A.H. & Gustafson, D.H., 1975. *Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes*, Glenview: Scott Foresman.
38. Okoli, C. & Pawlowski, S.D., 2004. The Delphi Method as a Research Tool: An Example, Design Considerations and Applications. *Information & Management*, 42(1), pp.15–29.
39. Hill, K.Q. & Fowles, J., 1975. The Methodological Worth of the Delphi Forecasting Technique. *Technological Forecasting and Social Change*, 7(2), pp.179–192.
40. Pousttchi, K., Moormann, J. & Felten, J., 2015. The Impact of New Media on Bank Processes: a Delphi Study. *International Journal of Electronic Business*, 12(1), p.1.
41. Argyrous, G., 2005. *Statistics for Research*, London: Sage.
42. Dajani, J.S., Sincoff, M.Z. & Talley, W.K., 1979. Stability and Agreement Criteria for the Termination of Delphi Studies. *Technological Forecasting and Social Change*, 13(1), pp.83–90.
43. Tobergte, D.R. & Curtis, S., 2013. Characteristics of the Turfgrass Industry in 2020: a Delphi Study with Implications for Agricultural Education Programs. *Journal of Chemical Information and Modeling*, 53(9), pp.1689–1699.
44. Putnam, J.W., Spiegel, A.N. & Bruininks, R.H., 1995. Future Directions in Education and Inclusion of Students with Disabilities: A Delphi Investigation. *Exceptional Children*, 61(6), pp.553–576.
45. English, J.M. & Kernan, G.L., 1976. The Prediction of Air Travel and Aircraft Technology to the Year 2000 Using the Delphi Method. *Transportation Research*, 10(1), pp.1–8.
46. No author, 2016. The DAO of Accrue. *The Economist*. Available at: http://www.economist.com/news/finance-and-economics/21699159-new-automated-investment-fund-has-attracted-stacks-digital-money-dao [Accessed August 9, 2016].
47. Wörner, D., von Bomhard, T., Schreier, Y.-P. & Bilgeri, D., 2016. The Bitcoin Ecosystem: Disruption Beyond Financial Services? In *ECIS 2016 Proceedings*.
48. Giaglis, G.M. & Kypriotaki, K., 2014. Towards an Agenda for Information Systems Research on Digital Currencies and Bitcoin. *Lecture Notes in Business Information Processing*, 183, pp.3–13.
49. Salmony, M., 2016. Blockchain – not for Payments? *BIT – Banking and Information Technology*, 56(2), pp.6–8.
50. Bharadwaj, A., El Sawy, O.A., Pavlou, P.A. & Venkatraman, N., 2013. Digital Business Strategy: Toward a Next Generation of Insights. *MIS Quarterly*, 37(2), pp.471–482.

From Shopping Aids to Fully Autonomous Mobile Self-checkouts – A Field Study in Retail

Denis Vučkovac[1], Pascal Fritzen[2], Klaus L. Fuchs[1], and Alexander Ilic[3]

[1] ETH Zurich, MTEC-IM, Zurich, Switzerland
{vdenis,klausfuchs}@ethz.ch
[2] TU Munich, CDTM, Munich, Germany
{pascal.fritzen}@cdtm.de
[3] University of St. Gallen, ITEM, St. Gallen, Switzerland
{alexander.ilic}@unisg.ch

Abstract. Self-checkout terminals allow integrating customers as active co-producers into a retailer's business processes. They have enjoyed increasing popularity in the past years since they allow saving costs and increasing customer satisfaction. Yet, they cannot be implemented in many retail settings, as the technology relies on retailer provided terminals and does not yet fully utilize the possibilities provided by mobile smartphones, which until recently have mostly served as decision or shopping aids. This paper presents steps towards and results from a field study of a purely mobile self-checkout solution that provides a more time efficient shopping experience to time-constrained users. We show that the time performance of app users is independent of store rush and that the time for a transaction is significantly lower for app users compared to regular shoppers during peak periods.

Keywords: shopping aids, mobile payments, self service technologies, retail, lost sales

1 Introduction

Recent implementations of self-service technologies (SST) such as self-checkout terminals enable users to scan and pay groceries without interaction with the store personal [1]. By skipping the waiting queues of regular registers, SST have been effective at increasing consumers' satisfaction and convenience [2, 3]. Self-checkout solutions today are usually implemented through large physical terminals that include barcode scanners and payment facilities [4]. Some solutions also provide the ability to scan items already during the shopping journey with dedicated handheld scanning devices or the user's mobile phones. However, even in these scenarios the digital shopping basket is transferred to a self-checkout terminal where the actual payment and integration of transaction data take place, creating new, unnecessary bottlenecks. As a consequence, capacity constraints remain during peak times [2]. While current self-checkout solutions are well suited for the digital transformation of regular grocery stores and have become well adopted among these [4], they are not easily applicable to

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Vuckovac, D.; Fritzen, P.; Fuchs, K.L.; Ilic, A. (2017): From Shopping Aids to Fully Autonomous Mobile Self-checkouts - A Field Study in Retail, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 927-941

a convenience store context. Unlike regular grocery stores, convenience formats typically have a high share of customers that only purchase very few items. In Europe, convenience stores are often located at train stations where space is highly constrained and expensive [5] and train schedules often generate significant peak loads of customers. During these high peak hours, convenience stores often experience a high share of lost sales as customers not willing or able to wait decide to abort their purchase or frequent competing stores [6, 7]. Due to the space requirements of traditional self-checkout solutions, it is often not feasible to dedicate the space needed to physical terminals to handle these peak loads more efficiently. While retailers have historically differentiated from competitors mainly through price, assortment and location [8], recent technologic advancements enable increasing differentiation through the use of interactive technologies [9]. We therefore develop a fully autonomous self-checkout solution suited for the fast-paced environment that combines self-scanning and mobile payments. We report findings from a digital transformation project including a field study conducted jointly with one of Europe's leading convenience store operators. Our contributions include a thorough analysis of the pre-transformation situation, critical aspects of the design and implementation of such an information system and results from an evaluation in a real-world pilot deployment with 46 buying users and 129 transactions in a period of 12 weeks. Our results provide valuable insights that are potentially applicable to other cases of digital transformation where end-users are involved in core business processes with own mobile devices. In addition to mitigating the lost sales problem, we open possibilities for forming a digital relationship between consumers and retailers.

The paper is structured as follows. We first study and summarize the relevant literature and illustrate our research context, questions and methodology. We illustrate the current situation and based on this we design a target shopping process adapted to our research context and provide an overview of our system architecture. We then discuss results from our field study and finish with a conclusion and outlook.

2 Related Work

2.1 Shopping aids

Many retailers have deployed technology solutions traditionally known as shopping or decision aids to assist shoppers in their shopping journey [11]. Common shopping aids, both online and in-store, can either help screen and narrow down available products or allow for the in-depth comparison of selected products [12]. While shopping aids in the form of in-store kiosks have been around in physical retail for a longer time [10], they have been used and studied much more in detail in online settings [11–13]. Online shopping aids have been shown to reduce search costs and increase convenience and the quality of purchase decisions [11, 12]. They can help enhance store loyalty but can also increase consumers' price and promotion sensitivity [13]. Furthermore, implementations of shopping aids in physical retailing besides in-store kiosks, such as smart shopping carts, can increase spending and re-patronage intentions as well as satisfaction and loyalty of users [14, 15]. Yet, such solutions require high up-front investments. However, the increasing ubiquity of smartphones paired with their mobile and personal nature, has made the operationalization of personalized shopping aids economically and technologically feasible in physical retailing on a much greater scale [16]. Thus, consumers are beginning to increasingly utilize mobile phones in the shopping process–currently mostly for information search in the pre-purchase phase and less for actual purchase transactions [17]. However, research has also shown that consumers shop differently online compared to offline [13] and that mobile shopping aids are overwhelming shoppers with information instead of providing key functionalities [18].

2.2 Mobile Payments

Mobile payments can be defined as all transactions in which consumers use a mobile phone to transfer money or funds from one party to another in exchange for goods or services [19–22]. In general, one can distinguish between *remote mobile payment applications*, in which transactions are made independently of the location of the user and *proximity mobile payment applications* in which a user's mobile phone communicates locally with the point-of-sales (PoS) [20]. While the former resemble online payment systems known from e-commerce settings, the latter require a solution to link the digital and physical economy [23]. Such solutions include utilization of short-range wireless communication protocols such as the Bluetooth protocol or near frequency communication (NFC) technology [23] or the scanning of a 2D barcode [24]. Mobile payments are frequently considered the "killer" application of mobile communication networks [25] and seen as the most critical driver for the success of mobile commerce [21]. Yet, the adoption and usage of mobile payments remains sobering and below expectations [20, 26–28]. Nevertheless, expectations are currently rising again due to increasing penetration of the NFC technology [26].

Several studies have highlighted the security and trust as prerequisites for the adoption of mobile payments [22, 27]. Research on mobile ticketing systems, where the transaction itself happens remotely and independently of the location, yet the ticket verification happens locally, show that convenience and speed have a big impact on perceived usefulness and the use of mobile ticketing systems [29, 30]. In such manner, people are willing to use even more complex ticketing systems when they are in a hurry or try to avoid queues [30]. Hence, the relative advantage of mobile payments is among others driven by possible queue avoidance and situational factors such as the presence of queues, support the adoption of mobile payment solutions [19]. Yet, current mobile payment solutions require mobile devices to interact with the same established PoS [20, 28] and thus require users to queue at the same register, offsetting potential advantages.

2.3 Self-service technologies

Self-service technologies (SST) can be defined as technological interfaces that allow customers to coproduce a service without employee interaction [1, 31–33]. Retailers mainly offer SST to reduce costs and improve customer experience [3]. The most frequently mentioned advantages of SST are convenience and speed [2]. Negative experiences include forced experiences of self-checkout uses, instances when self-checkout terminals have been closed at certain times of the day (i.e. night) and the fact that these terminals happen to be slower when there is a queue [2]. In accordance with this, researchers find complementary reasons for the use of self-checkout terminals and traditional checkouts at traditional retailers and advice practitioners to offer both in addition to another [34]: With respect to this, researchers have found evidence that SST customers tend to use self-checkout terminals for smaller baskets and might avoid items that require additional steps in the checkout process (e.g. fruits, vegetables etc.) [32]. In fact, small shopping baskets turn out to be a key reason for self-checkout usage and are only topped by long lines at traditional PoS [35].

The service quality of SST is mainly determined by functionality, enjoyment, design, assurance and convenience and has a positive impact on customer loyalty through customer satisfaction [3]. However, while the first trial and adoption of SST constitutes the most prominent obstacle since it usually involves significant behavior change [36], actual waiting time at regular checkout terminals acts as an important determinant for actual use [37]. From an industry perspective the main concern regarding self-checkout solutions is the potential increased risk of theft. While there is little evidence regarding the impact on retail shrinkage, the studies available suggest that there is no increase [38, 39]. To the best of our knowledge, there exists no implementation or analysis of a fully mobile self-checkout solution in grocery retailing.

3 Research Framework

3.1 Research Context

Our study is carried out in collaboration with a leading European convenience store retailer with a physical store network of over 1000 stores–most of them located at public transport transit points such as train stations. For a pilot study, we have selected three stores at Zurich main station. None of the stores currently have a self-checkout solution in place. In fact, traditional self-checkout solutions have been evaluated as unprofitable due to high place requirements and investment needs. Also, there is no mobile application or shopping aid offered by the retailer at the time of the pilot start and mobile payment support was introduced just shortly before the pilot start. Due to the fast paced nature of the business at the pilot store locations, there is no loyalty card system in place either. The retailer suspects and has anecdotal evidence that a high share of sales is lost due to long queues and time pressure of customers during peak hours.

3.2 Research Question and Methodology

The key challenge is to enable a self-checkout system that allows consumers to fully autonomously purchase products in a grocery store even under time pressure. To address this complex problem, we follow a design science approach [40]. First, we aim to leverage insights from store workers and managers as well as quantitative data from the three existing stores to get a better understanding of the extent and nature of the problem. Second, we translate the learnings in an iterative system design and implement a corresponding artefact consisting of a mobile application, an instore feedback component, and a backend system. Our development process focuses on previously outlined constructs from SST and mobile payments research, namely functionality, relative advantage, complexity and convenience [3, 36] and perceived risk, assurance, security and trust [3, 22, 27, 36]. We optimize these constructs over multiple test phases and iterations with various users. Finally, we evaluate the artifact in a 12 week pilot deployment. The corresponding research question we hereby aim to address is whether such an implemented artifact provides added benefit to users. We operationalize this by ultimately measuring its repeated use [41] and quantifying time performance during more and less busy periods of the day [42], as well as the difference to the regular shopping time during busy periods, as waiting time has been revealed to be a key issue for our industry partner and was shown to impact the satisfaction-loyalty relationship of retailers [6, 7].

4 Design and Evaluation

4.1 Analysis of Status Quo

The following analysis is based on one year transaction data for two stores and half a year for one of the three pilot stores with more than 1.2 Million receipts in total. The highest frequented store has on average about 2,300 transactions per day, followed by 1,500 transactions by the second and about 700 from the smallest of the three stores.

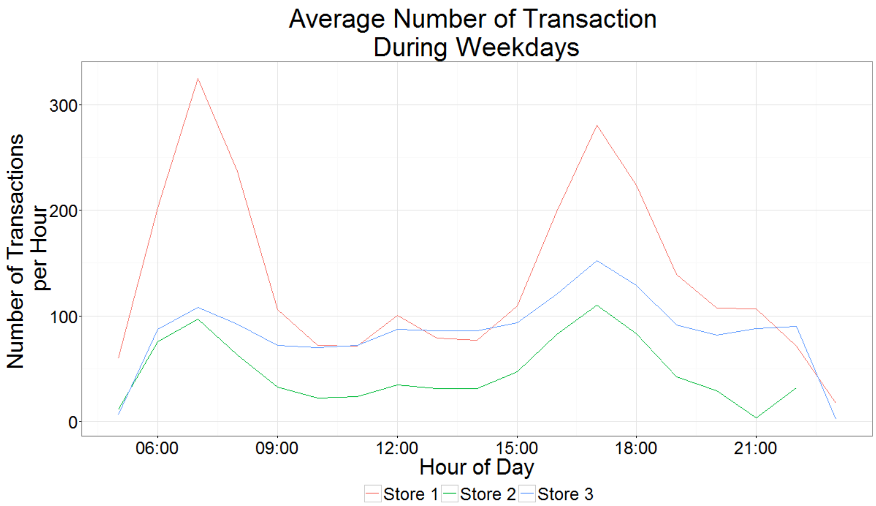


Figure 1. Distribution of average number of transactions per hour throughout weekdays

All of the three stores are characterized by very small shopping baskets with on average 1.5 to 1.8 items per transaction and an average basket value between 6 and 10 CHF. None of the stores currently have a self-checkout solution in place. The share of baskets that include alcohol, tobacco or services (ultimately, products that cannot be purchased through the mobile application), lies between 18 and 36%. The most popular products include soft drinks, beer, (self-service) coffee, tobacco, newspapers, magazines and bread and pastries. Our research partner does not offer a loyalty program and has thus no clear insight into consumption patterns of consumers. Due to the central location of our pilot stores and the high share of commuters, two of the three stores are characterized by high peak demand before and after working hours. Figure 1 illustrates the average daily pattern during weekdays for the three different stores. All of the three stores have their highest number of transactions either between 7 a.m. and 9 a.m. or between 5 p.m. and 7 p.m. Store 1, also the biggest in size, is characterized by the highest increase during peak hours of more than three times the average demand during. Especially this store exhibits very long queues during rush hours and customers are faced with long waiting times. We therefore argue that it is critical that a mobile self-checkout application minimizes the required time and effort for customers even during peak hours in the presence of long queues. Furthermore, it is important to mention that all the stores in our pilot are part of a chain and thus not stand alone stores, yet the mobile self-checkout application is only supported in the respective pilot stores.

4.2 Target shopping process

The key functionality of our mobile application is to buy physical products from a brick-and-mortar retailer without interaction with the PoS. The shopping process is composed of four main operations: A user has to check into the respective store before choosing products in order to see the correct price information. Users then select products by scanning the barcode printed on the product or shelf. After scanning the barcode, the user can adjust the desired quantity and then add the product to his (virtual) shopping basket. The user can then review his basket and start the checkout and payment process. The process is completed once the user confirms his transaction and the due amount is charged to the credit card information stored on the user's mobile device (Figure 2).

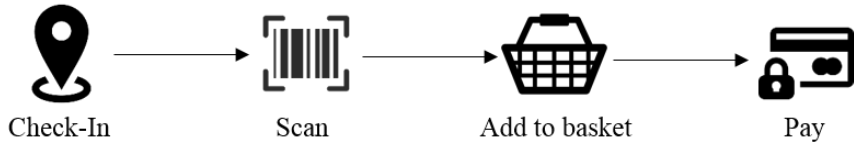


Figure 2. Illustration of target shopping process

The system was developed in several design iterations. The first iteration was conducted together with store workers and managers in order to come up with a first initial implementation which also allows store workers to track and identify valid transactions of users and thus emerged as a key requirement. The second design iteration took place in a lab setting where students used a first proof of concept and were faced with a typical shopping assignment. Based on feedback from seven testers, the design was refined. Eventually, it emerged that the testers favored a version of the self-checkout application that included a physical touch point (a QR code) within the retail store. The third design iteration took place in a real store setting where eight testers used the mobile self-checkout application executing real transactions in order to buy a product of their choice in one of the pilot stores. We again collected user feedback and use this to improve the mobile application's usability, mainly focusing on in app communication and the in-store feedback system.

4.3 System Design

Our system illustrated in Figure 3 consists of three components, the backend, an iOS application which allows users to self-scan and pay products with their smartphone and an in-store information system which provides feedback for successful transactions.

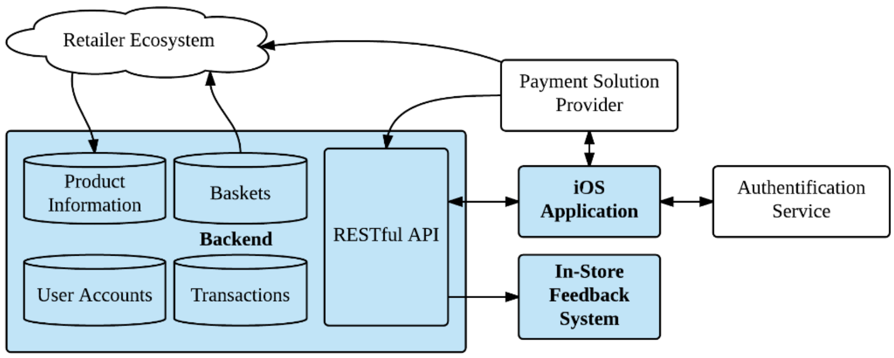


Figure 3. Overview of system architecture, developed artefact marked blue

The backend is responsible for storing all the information and communicating with the two clients. As our research partner's infrastructure relies highly on information systems that currently do not support application programming interfaces (APIs) that would allow the mobile application and the in-store feedback system to retrieve and store information, we mirror the retailer's product database and build our own backend. Product information is exported daily from our industry partner in order to account for changing prices. By predefining a specific data format we ensure a frictionless and correct re-integration of all transactions back into the retailer's information systems. The mobile self-checkout application "Scan&Go" is implemented for iOS systems and can be downloaded freely from the Apple App Store. The app features a short optional tutorial that outlines the steps to be performed in order to purchase a product. After registering with a phone number, a user has to fill out a short screening survey (six questions). Among others, users are asked about their motivation, demographics and average train station frequency. Since the service provided by our mobile application is highly location based, only at least multiple times per month at Zurich train station can participate in our pilot study[1]. Once registered and screened, a customer can use the mobile self-checkout application–if eligible.

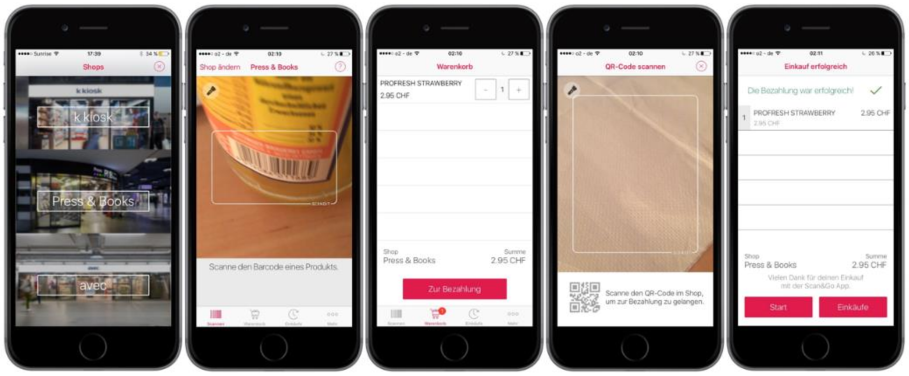


Figure 4. Screenshots of in-app view of 1) check-in, 2) barcode scanning, 3) virtual shopping basket summary, 4) QR-code scan and 5) confirmation and digital receipt

The in-store feedback system serves two purposes, theft prevention and public information signaling for successful purchases. It consists of a tablet that is mounted to a chipboard and features a poster with the QR which the user has to scan to confirm his payment. Although from a technologic perspective not required, the QR code serves three purposes: 1) it ensures that the user is in a pilot store, 2) it ensures that the user is in the right pilot store and did not accidentally check-into a wrong one and 3) guarantees that the user is located right in front of the tablet and store workers and other shoppers are able to match transaction feedback with a user[2]. The tablet constantly checks against the backend's API whether a transaction was successfully executed and approved by the payment service provider. In this case, the tablet, which is also visible from further away, turns green, features a confirmatory message and plays a positive, *cashregisterlike* sound. Based on user feedback, we believe, that a well designed and implemented in-store feedback system is key to support adoption and usage of such a purely mobile self-checkout application.

1 Since this train station is relatively big and features many stores of the same banner, even on several levels, this allows us to confirm a user's eligibility only to some extent.

2 While the first two purposes can be replaced with beacon technology the last cannot without compromising easy matching and/ or usability.

5 Results

5.1 User Recruitment

Users are recruited through three distinct channels. First, we email about 60 colleagues and friends. Furthermore, we distribute flyers to current customers either through the cashier or a dedicated student. Finally, we launch a small Facebook campaign targeted at people at the train station during the illustrated commuting hours.

5.2 User Statistics

The evaluation study ran in all three stores during 12 weeks from July to September 2016. We recruit 200 eligible users of our mobile application (out of 229 total registrations). 29 registered users were not eligible for participation either due to no supported payment method (11), age (11) and/ or a too low train station frequency (20). Users involved in the development of the mobile application are already excluded from this sample. 81% of all registered users are male and the majority is between 25 and 34 years old (35%) followed by 35 to 44 years (27%) and 18 to 24 years (22%). The majority (45%) of registered users states curiosity as their main motivation, followed by 38% that state time saving as the most important driver behind trying the mobile self-checkout application (convenience, 11%). 39% of all users are multiple times per week at the train station while slightly less, 36% of all registered users, visit the train station daily. 13% each state to be once a week or multiple times per month at the train station. In order to illustrate the consequences of long queues and waiting times during peak hours and quantify the previously hypothesized lost sales problem, we additionally ask users how often on average they have aborted or not even started a purchase due to long waiting times. In total, almost half of our sample state that this happens on average at least once a week, with about 10% claiming that this is even a daily problem. These statistics support our initial hypothesis and anecdotal evidence from store workers and managers. Although there is obviously self-selection involved in the acquisition of users for such a mobile self-checkout application, we conclude that there is great benefit for users and opportunity for retailers to differentiate – with a potentially highly lucrative business case for retailers able to address this issue by providing a faster shopping experience.

5.3 Conversion and Usage

Out of 200 eligible users 83 have scanned at least one product, either playing around with random products of other retailers or testing the app’s functionality with actual products from our research partner. The detailed user funnel is illustrated in Figure 5. No financial incentive for using the app has been given to users. In total 46 users (23% of all eligible users) have carried out at least one transaction through the app and in total 26 users have used the self-checkout application more than once for a product purchase. Thus, about 56% of all buying users have made more than one transaction.

While these numbers itself are above industry benchmarks of similar e- and m-commerce implementations [43, 44], conversion and usage numbers are even greater when accounting for the recruitment channel of customers. We distinguish between proximity recruitment (in-store through flyers) or distance recruitment (Email to colleagues and friends and Facebook campaign). The cohort type can be assigned to users based on the date of their respective registration. All of the relative conversion numbers of proximity recruited users are higher than for distance recruited ones.

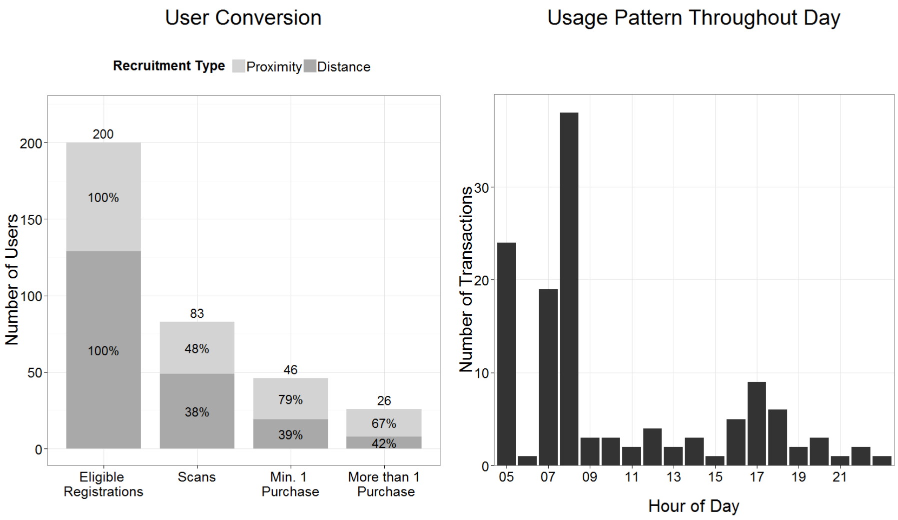


Figure 5. User conversion with number of users per action for proximity (flyer) and distance (Email, Facebook) recruited users (percentage numbers illustrating share of users relative to previous action for same recruitment type) and number of total transactions per hour

Out of all proximity recruited users, 48% have scanned an item using the app and 79% of these have ultimately also purchased something. In fact, about 38% of all proximity recruited users have made a purchase using the app, compared to 15% in the distance recruited group. Even more impressively, 46% of proximity recruited users stating "time saving" as their main motivation have made a purchase, and if stating to be at the train station "daily", the share of users making a purchase with the app is even 57% - with all of these (8) making at least one more purchase through the app. We report a total of 129 transactions issued through the mobile application with about 85% of all transactions made in Store 1. The most active user has a total number of transactions of 23, followed by one user with 16 and one with eight transactions. Figure 5 shows that most of the transactions through the app are made in the morning hours, specifically between 8 a.m. and 9 a.m. Compared to the overall sample, relative app usage is even greater during the morning peak.

5.4 Time Performance

In order to analyze the impact of increasing store rush and queues on the mobile app users, we measure the time required from 1) opening the app and 2) a first product scan to a successful transaction and its respective distribution throughout time and during peak and non-peak hours for all of the 129 transactions. For the first metric, 12 transactions (4 morning peak, 6 nonpeak, 2 afternoon peak) had to be excluded, as an instance of opening the app was either not reported or the app was obviously already opened much prior to entering the store (more than 10 minutes before the purchase). We find that the mean purchase time from app opening is about 99 seconds (st. deviation 86 seconds) with a median of 64 seconds. While our first metric has a relatively wide distribution, the purchasing time from a first scan is on average 40 seconds (st. deviation 38 seconds) with a median of 28 seconds and is thus more concentrated.

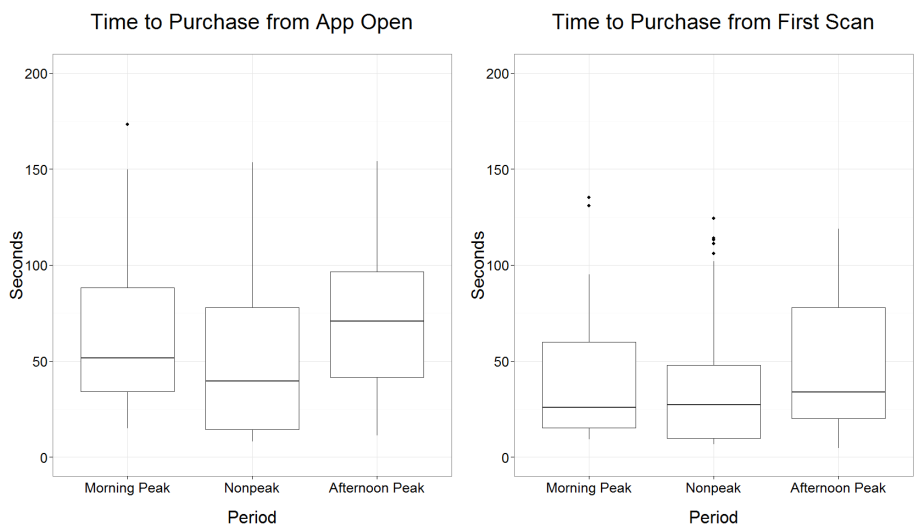


Figure 6. Standard boxplots (1st, 2nd and 3rd Quartile) of time to purchase from app open and first scan during peak (morning and afternoon) and nonpeak periods ($N_{morning\ peak}$ = 57, $N_{nonpeak}$ = 57, $N_{afternoon\ peak}$ = 15)

In order to compare our two measures throughout the day, we categorize the transactions into three different time intervals according to regular transaction data from the three pilot stores, with a morning peak period from 7 a.m. to 9 a.m., an afternoon peak period from 5 p.m. to 7 p.m. and a remaining nonpeak period. Our results in Figure 6 show almost equal or equal means and medians for both measures between the different periods of the day. We hypothesize that the somewhat higher median shopping time during the afternoon and evening period is due to less time pressure and more of a “strolling” shopping behavior of users. We attribute the wider distribution of our first metric to the differences in time needed to buy certain products (i.e. coffee from a self-service coffee machine – users could scan the barcodes before, after or while making the coffee) and also to the fact that some people already open the app before entering the target store. Due to the noisier distribution of the first metric, we argue that the purchase time from first scan is a more accurate measure of the actual shopping time required, as products bought have to be scanned in store whereas an instance of opening the app can occur outside of the store. In order to evaluate the performance of app users we derive a baseline by taking the shopping time for all regular shoppers during the morning peak period and compare this to app users during the same morning hours. For each users we measure the time from picking a first product from the shelf as this should correspond to a first scan of a product within the app. We collect a sample of 95 observations. The mean purchasing time for regular shoppers is 99 seconds (st. deviation 46 seconds) compared to 39 seconds (st. deviation 31 seconds) for app users during the morning peak.

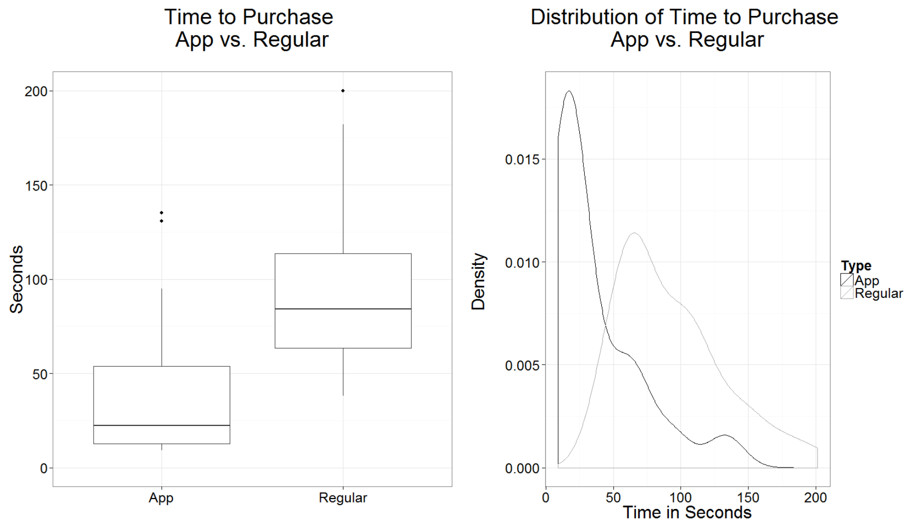


Figure 7. Standard boxplots (1st, 2nd and 3rd Quartile) and kernel density plots of purchase times for app users and regular shoppers during peak hours (N_{App} = 57, $N_{Regular}$ = 95)

Figure 7 shows the distribution of purchasing times for both app users and regular shoppers during the morning peak. We conduct a Wilcoxon-Mann-Whitney-Test [45] for non-normally distributed data and find statistically significant evidence that the time to purchase is smaller for app users (p = 8.56e-12).

6 Conclusion and Outlook

We have designed and implemented a mobile self-checkout application and tested its acceptance and usage in a first pilot at three stores at the central train station in Zurich. Learnings from the usage logs of 46 purchasing users, of which 26 are repeated purchasing users, illustrate positive value and consumer acceptance from such an application. An onboarding survey has provided additional insights into the demographics and motivation of our study participants and illustrated that almost half of all users regularly (at least once a week) fail to make purchases because of time pressure and long queues. With respect to this, we compare with the baseline time performance of regular shoppers during peak hours and show that app users on average save 60 seconds. Furthermore, we are able to show that the required purchasing time for app users is stable throughout the day, even in the presence of queues during morning and afternoon rush hours. Our study is a first step in understanding how to design a fully autonomous mobile self-checkout solution and transition from stationary to mobile retailing. In order to gain more insights into app adoption and usage and learn more about general consumption patterns of mobile app users, we aim to extend our study further with more participants over a longer time period and collect additional data on satisfaction of mobile app users. Furthermore, besides technical obstacles one obvious reason for retailers not implementing a purely mobile self-checkout solution is easier control of actual product purchases when using self-checkout terminals. Consequently, we aim to measure changes in inventory shrinkage at the end of our extended study in order to evaluate the impact of such an application on theft and inventory shrinkage. A self-checkout application as presented in our study allows retailers to offer valuable information systems based services to consumers beyond only the provisioning of additional product information and allows retailers to pursue new digital business models. It can help retailers differentiate from competitors in an increasingly crowded market place. Thereby, consumers profit from a faster and more convenient shopping experience. Furthermore, in the absence of existing loyalty programs, such a mobile app provides more sophisticated insights into individual shopping patterns.

References

1. Kallweit, K., Spreer, P., Toporowski, W.: Why do customers use self-service information technologies in retail ? The mediating effect of perceived service quality. Journal of Retailing and Consumer Services. 21, 268–276 (2014).
2. Anitsal, I., Daniel J. Flint: Exploring Customers’ Perceptions in Creating and Delivering Value. Services Marketing Quarterly. 27, 57–72 (2006).
3. Demirci Orel, F., Kara, A.: Supermarket self-checkout service quality, customer satisfaction, and loyalty: Empirical evidence from an emerging market. Journal of Retailing and Consumer Services. 21, 118–129 (2014).
4. NCR: Self-Checkout: a Global Consumer Perspective. (2014).
5. Debrezion, G., Pels, E., Rietveld, P.: The impact of railway stations on residential and commercial property value: A meta-analysis. Journal of Real Estate Finance and Economics. 35, 161–180 (2007).
6. Tom, G., Lucey, S.: Waiting time delays and customer satisfaction in supermarkets. Journal of Services Marketing. 9, 20–29 (1995).
7. Bielen, F., Demoulin, N.: Waiting time influence on the satisfaction-loyalty relationship in services. Managing Service Quality. 17, 174–193 (2007).
8. Krafft, M., Mantrala, M.K.: Retailing in the 21st Century: Current and Future Trends. (2010).

9. Varadarajan, R., Srinivasan, R., Vadakkepatt, G.G., Yadav, M.S., Pavlou, P.A., Krishnamurthy, S., Krause, T.: Interactive technologies and retailing strategy: A review, conceptual framework and future research directions. Journal of Interactive Marketing. 24, 96–110 (2010).
10. Rowley, J., Slack, F.: Kiosks in retailing: the quiet revolution. International Journal of Retail & Distribution Management. 31, 329–339 (2003).
11. Vijayasarathy, L.R., Jones, J.M.: Do Internet Shopping Aids Make a Difference? An Empirical Investigation. Electronic Markets. 11, 75–83 (2001).
12. Haubl, G., Trifts, V.: Consumer Decision Making Decision Aids in Online Effects of Shopping The Environments. Marketing Science. 19, 4–21 (2000).
13. Wei Shi, S., Zhang, J.: Usage Experience with Decision Aids and Evolution of Online Purchase Behavior. Marketing Science. 33, 871–882 (2014).
14. Ittersum, K. van, Wansink, B., Pennings, J.M.E., Sheehan, D.: Smart Shopping Carts: How Real-Time Feedback Finluences Shopping. 77, 21–36 (2013).
15. Lee, H.-J.: Consumer-to-store employee and consumer-to-self-service technology (SST) interactions in a retail setting. International Journal of Retail & Distribution Management. 43, 676–692 (2015).
16. Shankar, V., Venkatesh, A., Hofacker, C., Naik, P.: Mobile marketing in the retailing environment: Current insights and future research avenues. Journal of Interactive Marketing. 24, 111–120 (2010).
17. Holmes, A., Byrne, A., Rowley, J.: Mobile shopping behaviour: insights into attitudes, shopping process involvement and location. International Journal of Retail & Distribution Management. 42, 25–39 (2014).
18. Kalnikaite, V., Bird, J., Rogers, Y.: Decision-making in the aisles: Informing, overwhelming or nudging supermarket shoppers? Personal and Ubiquitous Computing. 17, 1247–1259 (2013).
19. Mallat, N.: Exploring consumer adoption of mobile payments - A qualitative study. Journal of Strategic Information Systems. 16, 413–432 (2007).
20. Chandra, S.: Evaluating the Role of Trust in Consumer Adoption of Mobile Payment Systems: An Empirical Analysis. Communications of the Association for Information Systems. 27, 561–588 (2010).
21. Yang, S., Lu, Y., Gupta, S., Cao, Y., Zhang, R.: Mobile payment services adoption across time: An empirical study of the effects of behavioral beliefs, social influences, and personal traits. Computers in Human Behavior. 28, 129–142 (2012).
22. Dahlberg, T., Mallat, N., Ondrus, J., Zmijewska, A.: Past, present and future of mobile payments research: A literature review. Electronic Commerce Research and Applications. 7, 165–181 (2008).
23. Vazquez-Briseno, M., Hirata, F.I., Sanches-Lopez, J. de D., Jimenez-Garcia, E., Navarro-Cota, C., Nieto-Hipolito, J.I.: Using RFID/NFC and QR-Code in Mobile Phones to Link the Physical and the Digital World. In: Deliyannis, I. and Smiljanic, T. (eds.) Interactive Multimedia. pp. 219–242. InTech (2012).
24. Gao, J., Kulkarni, V., Ranavat, H., Chang, L., Mei, H.: A 2D barcode-based mobile payment system. 3rd International Conference on Multimedia and Ubiquitous Engineering, MUE 2009. 320–329 (2009).
25. Liebana-Cabanillas, F., Sanchez-Fernandez, J., Munoz-Leiva, F.: Antecedents of the adoption of the new mobile payment systems: The moderating effect of age. Computers in Human Behavior. 35, 464–478 (2014).

26. De Reuver, M., Verschuur, E., Nikayin, F., Cerpa, N., Bouwman, H.: Collective action for mobile payment platforms: A case study on collaboration issues between banks and telecom operators. Electronic Commerce Research and Applications. 14, 331–344 (2015).
27. Dahlberg, T., Guo, J., Ondrus, J.: A critical review of mobile payment research. Electronic Commerce Research and Applications. 14, 265–284 (2015).
28. Slade, E., Williams, M., Dwivedi, Y., Piercy, N.: Exploring consumer adoption of proximity mobile payments. Journal of Strategic Marketing. 4488, 1–15 (2014).
29. Cheng, Y.H., Huang, T.Y.: High speed rail passengers' mobile ticketing adoption. Transportation Research Part C: Emerging Technologies. 30, 143–160 (2013).
30. Mallat, N., Rossi, M., Tuunainen, V.K., Öörni, A.: An empirical investigation of mobile ticketing service adoption in public transportation. Personal and Ubiquitous Computing. 12, 57–65 (2008).
31. Curran, J.M., Meuter, M.L.: Self-service technology adoption: comparing three technologies. Journal of Services Marketing. 19, 103–113 (2005).
32. Wang, C., Harris, J., Patterson, P.G.: Customer choice of self-service technology: the roles of situational influences and past experience. Journal of Service Management. 23, 54–78 (2012).
33. Eastlick, M.A., Ratto, C., Lotz, S.L., Mishra, A.: Exploring antecedents of attitude toward co-producing a retail checkout service utilizing a self-service technology. The International Review of Retail, Distribution and Consumer Research. 22, 337–364 (2012).
34. Dabholkar, P.A., L. Michelle Bobbitt, Eun-Ju Lee: Understanding consumer motivation and behavior related to self-scanning in retailing: Implications for strategy and research on technology-based self-service. Internation Journal of Service Industry Management. 14, 59–95 (2003).
35. Smith, A.D.: Exploring the inherent benefits of RFID and automated self-serve checkouts in a B2C environment. International Journal of Business Information Systems. 1, 149 (2005).
36. Meuter, M.L., Bitner, M.J., Ostrom, A.L., Brown, S.W.: Choosing Among Alternative Service Delivery Modes: An Investigation of Customer Trial of Self-Service Technologies. Journal of Marketing. 69, 61–83 (2005).
37. Penttinen, E., Kahila, T.-R., Rönkkö, M., Saarinen, T.: Triggering Intention to Use to Actual Use - Empirical Evidence from Self-Service Checkout (SCO) Systems. In: Proceedings of the Annual Hawaii International Conference on System Sciences. pp. 3347–3355 (2014).
38. Beck, A.: Self-scan checkouts and retail loss: Understanding the risk and minimising the threat. Security Journal. 24, 199–215 (2011).
39. ECR: The Impact and Control of Shrinkage at Self Scan Checkouts. (2011).
40. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly. 28, 75–105 (2004).
41. Brown, S. a., Venkatesh, V., Goyal, S.: Expectation Confirmation in Technology Use. Information Systems Research. 23, 474–487 (2012).
42. Andriulo, S., Salento, U., Salento, U., Salento, U.: Mobile self-checkout systems in the FMCG retail sector: A comparison analysis. International Journal of RF Technologies Research and Applications. (2014).
43. Hasan, L., Morris, A., Probets, S.: Using google analytics to evaluate the usability of e-commerce sites. In: Human Centered Design. pp. 697–706 (2009).
44. Criteo: State of Mobile Commerce. 41 (2015).
45. Mann, H.B., Whitney, D.R.: On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics. 18, 50–60 (1947).

Ideate. Collaborate. Repeat. A Research Agenda for Idea Generation, Collaboration and Evaluation in Open Innovation

Thomas Wagenknecht[1], Jan Crommelinck[1], Timm Teubner[2], and Christof Weinhardt[2]

[1] FZI Research Center for Information Technology, Berlin, Germany
{wagenknecht,crommelinck}@fzi.de
[2] Karlsruhe Institute of Technology, Karlsruhe, Germany
{timm.teubner,christof.weinhardt}@kit.edu

Abstract. Open innovation has been and remains to be a rapidly changing field of research in Information Systems and various other disciplines. With the rise of professional open innovation platforms and the emergence of crowdsourcing as well as employee-driven innovation, studies on the front-end of open innovation – namely idea generation, collaboration and evaluation – are facing new challenges. In this structured literature review, we analyze a large body of prior research in order to derive a framework, which is able to classify and reflect the lively debate on open innovation. In addition, we identify important implications for practitioners with advise on the design of open innovation systems. Moreover, our study identifies several promising areas for future research.

Keywords: open innovation, literature review, generation, collaboration, evaluation

1 Introduction

More than a decade after its conceptual inception by Chesbrough [12], open innovation (OI) still receives remarkable attention by scholars. It developed into an established research field in Technology and Innovation Management as well as Information Systems [17]. Many organizations, including public and corporate agents, have established OI platforms to solicit innovative ideas from a broad base of users. OI is an important means to create disruptive business innovations, rapidly changing existing and shaping new business models, processes and products [2]. For instance, Dell's ongoing "IdeaStorm" generated more than 20,000 suggestions for product improvements from thousands of registered users [6]. Because of such vast numbers of participants and proposals, an OI contest is likely to produce superior ideas and solutions that are able to compete with experts and innovators from corporate research and development (R&D) units – a proposition in line with the "wisdom of the crowds" theory [2, 31, 45, 47]. However, previous research suggests that these large idea collections in OI processes also tend to produce a number of highly redundant ideas and suggestions that greatly vary in terms of quality [8, 45, 46]. While about a third of

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Wagenknecht, T.; Crommelinck, J.; Tebner, T.; Weinhardt, C. (2017): Ideate. Collaborate. Repeat.
A Research Agenda for Idea Generation, Collaboration and Evaluation in Open Innovation,
in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung
Wirtschaftsinformatik (WI 2017), St. Gallen, S. 942-956

the ideas might be great, a majority is either mediocre or of low quality and is hence discarded as scrap [46, 28]. This is one reason why organizations often refrain from having their own experts evaluate each proposal, and resort to ask all users in an OI engagement to collectively evaluate and develop ideas further.

The environment for OI platforms is rapidly changing. At the beginning, many firms tried to set up proprietary OI systems, but with the rise of crowdsourcing [23] and professional OI platform providers (e.g., Hyve, Exago), OI might turn into a common form of R&D in leading corporations, as well as being more easily accessible to small and medium-sized enterprises (SMEs) and even individuals [17]. Moreover, with the emergence of topics such as employee-driven innovation [21, 13] and computer-supported organizational participation [57], OI might face new requirements within firms. Moreover, while process facilitators of OI engagements were able to profit from its novel character for a long time, they might face the problem of engaging less technology-savvy users and keep users engaged and active over a longer period of time going forward – both inside and outside the company. Considering the recent surge in research publications on OI and the changing environment for it, there is a need for a unified and structured framework that is able to both classify prior studies and guide future research.

In this paper, we therefore explore how extant research has analyzed the determinants for idea generation, collaboration and evaluation in OI – representing the key elements of the OI front-end [24] – to derive lessons on how OI systems need to be designed in the future in order to produce innovative solutions. In effect, we can also highlight areas for future research.

To do so, we analyzed 50 articles identified by means of a structured literature review [60]. In order to support researchers and practitioners in identifying well-studied and under-researched areas, we provide a concept matrix and a related framework that illustrate the flow of typical OI processes with the most relevant components. Both classify and summarize the studies along the sphere and sources the OI system addresses, the type of IT artifact, as well as the subject, testable propositions and methodology of the research. Our findings illustrate that idea generation and idea evaluation were almost equally often considered by the literature, mostly analyzing the collaboration processes. We find that researchers often recommend interactive ideation processes to increase the proposals' potential. Moreover, using multi-attributive rating scales was regularly found to strengthen decision quality. We criticize that many researchers developed OI systems on their own, rather than adopting prior development. Finally, we highlight areas for future research, including researching phenomena such as information cascades as well as the difference in idea generation and evaluation by internal versus external crowds.

In what follows, we explore the background and set the boundaries of our research in Section 2. We then introduce our methodology in Section 3 and describe the literature review process in detail. In Section 4, we report our results and discuss implications for practitioners as well as future research in Section 5. Section 6 draws a conclusion.

2 Background

For more than a decade, research on OI has been and continues to be a rapidly emerging and developing field of study in Information Systems and various other disciplines, such as Economics and Management Science [24, 2, 61]. Unsurprisingly, scholars have proposed a number of definitions and models that aim to describe OI. For the purpose of this literature review, we use a broad approach by Chesbrough [12], who defined OI as the use of purposeful inflows and outflows of knowledge to stimulate internal innovation, and expand the markets for the external use of innovation.

In practice, OI is often implemented using an idea contest. Adamcyzk et al. [2] referred to this as IT-based and time-limited competitions by organizations calling on the general public or a specific target group to propose innovative solutions. In doing so, the organizers make use of the expertise, skills and creativity of a crowd of users.

Hrastinski et al. [24] classified OI systems as technologies for idea management, problem solving and innovation (analysis). Their front-end typically comprises features and processes that support users in generating proposals and, consequently, developing and evaluating them. The systems might include sophisticated measurement tools to enable the evaluation process. Users might also be incentivized by rewards and recognition to participate in OI engagements. We will refer to these two common functionalities of OI systems as idea generation and evaluation, which can both happen with or without collaboration amongst users of the system [24].

Usually OI systems are implemented in the public sphere on one hand, by actors such as governments and non-governmental organizations. On the other hand, OI systems are especially popular with firms. Moreover, Gassmann et al. [17] suggested that universities and other academic organizations are engaging in OI too. These three broad spheres already hint at the possible target groups of OI processes. These are typically crowds that are either internal to the facilitating organization (e.g., employees, members) or external (e.g., customers, general public). In addition, OI facilitators often involve an (independent) expert committee to evaluate user-generated content [2]. Thus, considering these three broad spheres and target groups, research on OI is able to investigate various factors and their effects. Reviewing studies in Economics and Management Science, Adamczyk et al. [2] suggested that scholars are mainly concerned with assessing (1) the quality of idea generation processes, (2) the efficient design of OI processes, as well as (3) the users' motives to participate in OI engagements.

3 Research Method

In what follows, we describe our method for data collection, which builds the basis for the subsequent analysis. First, in order to provide a clear scope for this literature review, we need to set the boundaries of research [60]. We focus on what Hratinski et al. [24] referred to as the front-end of OI systems; that is, studies on computer-supported tools for the generation and evaluation of creative and valuable ideas and solutions in OI, including their collaborative development and rating. Thus, we consider the process from the point at which a facilitator opted to use an OI system until the point at which it comes to the decision of whether and how an idea shall be implemented. Hratinski et al. [24] referred to the latter as the back-end of OI. Also, research on new product development that does not explicitly refer to an OI process (e.g., by using data from an OI platform) is hence beyond the scope of our study. Furthermore, as we expect to find a large number of research articles on OI, we need to focus on studies that contribute most to the cumulative building of knowledge in the Information Systems literature by providing an advance to previous propositions and models [20]. Thus, we only include research that proposes the design of a solution for a pre-defined problem along with some form of demonstration and/or evaluation [43, 20].

3.1 Data Collection

Following the principles of Webster and Watson [60], we conducted an in-depth topic-based literature review focusing on idea generation, collaboration and evaluation in OI systems.

As OI represents an interdisciplinary and emerging research field, we included all relevant research published in journals listed in the ABS Academic Journal Guide 2015 [1] in the research subject areas of (1) Economics, Econometrics and Statistics, (2) Information Management, (3) Marketing, (4) Innovation and (5) Operations Research and Management. As we focus on Information Systems in particular, we also included full papers published in the seven leading generic Information Systems conference proceedings as recognized by ACPHIS [3]. To this list of conferences, we added CHI as the leading conference on Human-Computer-Interaction [67].

To investigate the literature base, we concentrated on the following databases: ProQuest (ABI/INFORM), Elsevier, IEEE, ACM, JSTOR, Web of Science, and EBSCOhost. Furthermore, AIS electronic library was accessed to review relevant conference proceedings.

For the research database search, we used a set of keyword combinations. In order to cover the broader literature on OI, we paired “open innovation” with “process”, “system”, “engagement” and “design”. Additionally, we combined “innovation” with “contest” and “tournaments” as these words are sometimes used as quasi-synonyms for OI engagements. Moreover, we wanted to cover more detailed studies on the sub-processes of OI activities. Therefore, we used a broad set of words we combined with “idea”, namely “generation”, “collaboration”, “evaluation”, along with a number of synonyms such as “assessment”, “voting”, “rating”, “ranking”, “screening”, and “filtering”, as well as “competition” and “management”.

Articles published before the year 2000 were excluded from our research, since computer-supported ideation and evaluation in OI was not properly defined in the last century.

Our literature search was conducted in three steps from April to May 2016. First, keyword search resulted in 212 articles being selected based on their title and abstract. We then removed duplicates and irrelevant articles. For instance, many articles investigated creativity or evaluation techniques in closed innovation environments. Other scholars analyzed managerial consequences or the implementation process of new ideas gained from OI, which is also beyond the scope of our study. Moreover, articles from publications other than those listed in the ABS Academic Journal Guide 2015 and conference proceedings recognized by ACPHIS were excluded to ensure a high level of quality. Second, the remaining 88 articles were analyzed in more detail focusing on their methodology and findings. Articles not satisfying the conditions set in our boundaries of research in Section 2 were excluded from our subsequent analysis. For instance, some studies implemented a system and refrained from evaluating it properly.

Articles satisfying the conditions introduced in Section 2 formed the basis of our third and last step. There, we conducted backward and forward searches, leading to 13 additional articles. In total, this structured review process resulted in a sample of 29 journal and 21 conference articles.

3.2 Data Analysis

Following Webster and Watson [60], we categorized the literature according to topic-related concepts as motivated in Section 2. First, we classified the articles based on the sphere that the study was conducted in, meaning public, corporate or academic [17]. Second, we extended our literature review by categorizing the type of source the research examined. On the basis of typical OI target groups, a source is either an external or internal crowd developing and/or evaluating ideas. Besides these crowds, an independent expert committee can also serve as a source of information [2]. Third, we also analyzed whether the research in our literature review proposed and evaluated an IT artifact of some sort. The definition of IT artifacts is subject to debate in the Information Systems literature [20]. Yet, we followed the definition by Peffers et al. [43], describing an artifact as something artificial, constructed by humans, which can be “any designed object in which a research contribution is embedded in the design” (p. 55). Furthermore, we adopted Gregor and Jones’ [20] classification of artifacts in terms of models, principles and methods. We also added the category of full system, which describes whether an artifact includes models, principles and methods to enable idea generation, collaboration and evaluation. Fourth, we categorized each study by its main research subject. As we focus on the front-end of OI as defined by Hratinski et al. [24], the three categories are idea generation, collaboration and evaluation. Moreover, each article investigated OI with regards to some form of testable proposition by introducing a quantitative, statistical analysis or through heuristic propositions [20, 56]. With regard to research on OI, we categorized the studies according to whether they (1) perform quality assessment, (2) analyze the efficiency of a process or (3) investigate user motivation [2]. Additionally, we analyzed in which sphere each study was conducted in. Finally, we categorized the identified literature according to the methodology used. Building on Palvia et al. [40], we limited these categories to frameworks/models, literature reviews, case studies, surveys, mathematical models and interviews. Two researchers classified the literature independently. Few inconsistencies were discussed and re-evaluated in order to reach a common understanding and resolve discrepancies.

4 Results

Our results point out that OI in general and idea generation, collaboration and evaluation in particular, recently received increased attention by researchers (see

Table 1). Most studies were published in conference proceedings, followed by research published in leading journals (see Table 2). Moreover, retrieving 19 articles in research fields such as Technology and Innovation Management as well as Marketing, confirms that OI is a constantly evolving, interdisciplinary field of research.

Table 1. Publications by time frame

Time frame	*2000-2003*	*2004-2007*	*2008-2011*	*2012-2016*
Publications	0	6	18	26

Table 2. Publications by research outlet

ABS Ranking [1]	*4/4**	*3*	*2*	*Conference*
Publications	15	8	6	21

We developed a concept matrix that categorizes each study (see Table 3). The concept matrix follows the outline of our data analysis in that it is structured in terms of the sphere and source, the type of IT artifact as well as the subject of the study, its testable propositions and methodology. In terms of idea generation we found 11 articles, compared to 16 articles investigating idea evaluation exclusively, whereas 21 articles covered both subjects at least partially. Interestingly, we found that researchers covered collaboration only in conjunction with either idea generation or evaluation, but never as a stand-alone research subject. Collaboration was investigated almost equally for generation and evaluation (39 vs. 33). This arises from the fact that many articles investigate OI systems that rely on collaboration.

With regard to the testable proposition, the vast majority of all articles covered at least some kind of quality assessment. In many cases, studies analyzed the quality of user-generated ideas through evaluations by experts committees [e.g., 46, 8, 28, 27, 32]. Thus, the propositions were both quantitative as well as heuristic in nature [56], as the experts used standardized rating methods to express their personal evaluation. Some studies took a more quantifiable approach, for instance, by evaluating the degree of user participation and activity (e.g., based on the number of executed trades or submitted ideas) on an idea market platform [51]. Two other studies conducted social network analyses [25, 7], which assessed both quality and quantity of user interactions. Moreover, 12 studies were concerned with evaluating efficiency of the processes of an OI engagement. Most often, this was the case for research on rating scales, where scholars tested how fast and accurate participants were able to conduct an evaluation task [e.g., 46, 8, 28, 14, 5]. Only one article (despite the

Table 3. Concept Matrix

	Sphere			Sources				Subject			Testable propositions			Methodology					
	Academic	Firm	Public	experts	internal crowds	external crowd	IT artifact	Generation	Collaboration	Evaluation	Quality Assessment	Efficiency	User Motivation	Framework/Model	Literature Review	Case Study	Survey	Math. Model	Interview
[2] Adamczyk et al 2012	◐	◐	◐	◐	◐	◐		●	●	●					●				
[4] Bailey and Horvitz 2010		●		●	●		FS	●	●	●	●		●			●			●
[5] Bao et al. 2011	●					●	ME			●	●		◐			●			
[6] Bayus 2013		●				●	FS	●	◐		●		◐			●			
[7] Björk and Magnusson 2009		●				●		◐	●	◐	●					●			●
[8] Blohm et al. 2010		●		◐		●	ME	●	●		◐		●			●			
[9] Blohm et al. 2011	●			◐		●	ME			●	●		●			●	●		
[10] Boudreau et al 2011		●	●			●		●	◐				●			●			
[11] Bullinger et al 2010			●			●		●	●		●		●			●	●		●
[13] Ciriello et al. 2016		●			●		ME			●	◐	◐		●					●
[14] Dean et al. 2006	●			◐		◐	PR			●	●			●					
[16] Feldmann et al. 2014	●	●			◐	●			●	●	●		●			●			
[19] Görs et al. 2012	●					●	ME			●	●	●	●			●		●	
[22] Horton et al. 2016	●	◐				●	PR			●	●					◐	◐	●	
[24] Hrastinski et al. 2010		●	●		●	●		●	●	●	◐	◐	◐	◐	●				
[25] Hutter et al. 2011		●				●		●	●	◐	●		◐			●	●		
[26] Jung et al. 2010	●					●	ME	●	●		◐		●			●			
[27] Kathan et al. 2015		●				●	FS	●	●		●		◐			●			
[28] Klein and Garcia 2015	●	●		◐	●		ME			●	●	●				●			
[29] Kornish and Ulrich 2011	●					●		●			●	●				●			
[30] Kristensson et al 2004		●		●	●	●		●			●					●			
[32] Lauto and Valentin 2016		●			●					●	●					●			
[33] Lee and Seo 2013		●				●			◐	●	●		●			●			
[34] Leimeister et al 2009		●		◐		●	FS	●	●	●	●		●	◐		●	●		
[35] Luo and Toubia 2015	●					●	PR	●		◐	●		●			●	◐		
[36] Magnusson et al 2014	●			●						●	●					●			
[37] Muller et al. 2013		●			●		FS	●	●	●	◐		●			●			
[38] Natalicchio et al 2014	●	●	●	●	●	●		●	●	●	●		●		●				
[41] Pashkina and Indulska 2011		●				●	MO	●			◐	◐	◐	●					
[42] Pedersen et al. 2013	◐	◐	◐	◐	◐	◐		◐	◐	◐	◐	◐	◐		●				
[44] Piller and Walcher 2006		●		◐		●	FS	◐		●	●					●			◐
[45] Poetz and Schreier 2014		◐	◐					●		◐	●					◐			
[46] Riedl et al. 2013	●	◐		◐		●	ME			●	●	●				●			
[47] Riedl et al. 2010	●	●		◐		●	ME			●	●		●			●	◐		
[48] Sawhney et al. 2005		●			●			●	●	◐	◐		◐			●			
[49] Scheiner 2015			●			●	FS	●	●	●			●			●	●		●
[50] Siemon et al. 2016	●					●	FS	●		◐	●		●			●	●		
[51] Soukhoroukova et al. 2012		●		◐	●		FS	●	●	●	●	●	◐			●			
[52] Stieglitz and Hassannia 2016		●			●	●		◐		●	●		●			●			
[53] Terwiesch and Xu 2008	●					●	PR	◐		◐	●		●					●	
[54] Toubia 2006	●					●	MO	●					●			●			
[55] Toubia and Flores 2007	●	◐				●	ME			●	●	◐				◐		●	
[58] Walter and Back 2013	●		●	●		●	ME			●	●	◐		◐		●			
[59] Walter and Back 2011		●	●			●			●	●	●		●	●		●			
[61] West and Bogers 2013	●	●	●			●		●	●	●	●		●		●				
[62] Wu and Fang 2010		●				●		●	●				●			●	●		
[63] Xu and Bailey 2012		●				●	MO			●	●					◐		◐	
[64] Yu and Nickerson 2011	●					●	ME	●	◐	◐	●					●			
[65] Yücesan 2013	●					●	MO			●		●						●	
[66] Zimmerling et al. 2016		●		◐	●		FS	◐	◐	●			●				●		●

ME: Method, FS: Full system, PR: Principle, MO: Model literature reviews) examined efficiency in collaborative idea generation; in this case by analyzing the redundancy of idea proposals [29]. Also user motivation was measured by many scholars, mostly by means of surveys or interviews. These studies asked for users' motives to participate and their satisfaction with the OI system. However, user motivation was often covered as an additional topic rather than being the main research question.

The majority of extant research investigated OI in the context of the sphere of the firm. Many papers also included an academic perspective, while only a fifth of the studies addressed the public sphere.

Many studies did not propose and evaluate an IT artifact. Surprisingly though, those that did often proposed a full OI system, which covered all the features described by Hratinski et al. [24] as the front-end of OI systems. Among them, many were studies in the domain of gamification, which focused on topics of user involvement through gamified reward systems and rankings to provide a gripping user experience [49, 66, 16, 54].

Most studies dealt with an external crowd as its source for idea generation, collaboration and/or evaluation. Other studies contributed to the overall trend of employee-driven innovation by sourcing ideas or evaluation from an internal crowd of employees [4, 13]. Moreover, some studies asked experts to assess the quality of user-

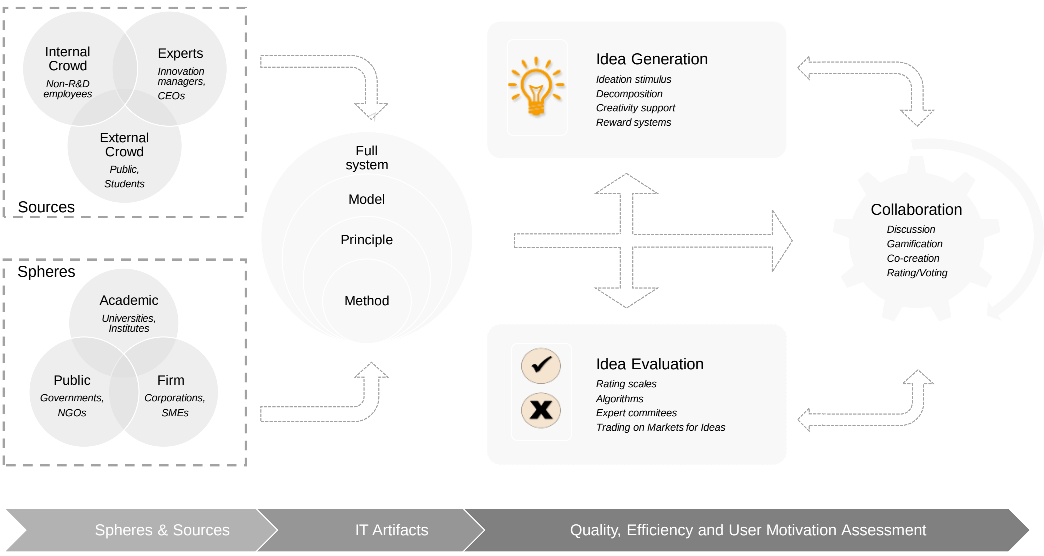


Figure 1. Framework for idea generation, collaboration and evaluation in OI

generated ideas or ratings. Magnusson et al. [36] was the only study to solely focus on experts as a source of information. The study analyzed different idea screening procedures by asking experts to rate ideas retrieved from an OI contest.

In terms of the methodologies, we find a rather clear picture. Despite the variety of research areas covered in our literature review, the vast majority of articles employed case studies [4, 5, 6, 7, 8, 9, 10, 11, 16, 19, 22, 25, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 44, 45, 46, 47, 48, 49, 50, 51, 52, 54, 55, 58, 59, 62, 63, 64]. Scholars either set up their own OI systems and applied them in practice – collaborating with firms, students or the general public – or extracted data from existing OI platforms (e.g., Dell's IdeaStorm [6] or Starbuck's MyStarbucksIdeas.com [4]). These case studies were sometimes supported by surveys [9, 11, 22, 25, 34, 35, 47, 49, 50, 62, 66] and/or interviews [4, 7, 11, 13, 44, 49, 66] in order to explore users' motives to engage in and perception of OI engagements. Seven studies developed frameworks and models of OI [13, 14, 24, 34, 41, 58, 59]. Moreover, six studies developed mathematical models in order to investigate the optimal design of OI processes [19, 22, 53, 55, 63, 65]. Furthermore, we found five literature reviews [2, 24, 38, 42, 61]. They were at least two years old and examined distinctively different research questions than our study. For instance, they examined literature on boundary areas of OI, such as markets for ideas [38] or crowdsourcing [42].

Investigating lessons for the design of OI systems, we find that idea generation and idea evaluation were almost equally often considered by the literature. Most studies did so by also analyzing the collaboration processes. Both Bullinger et al. [11] and Blohm et al. [8] suggest that collaborative ideation outperforms non-collaborative approaches. Moreover, research [18] established that the point in time when users are involved in collaborative processes is crucial. Moreover, Luo and Toubia [35] also emphasized that decomposing an idea and providing stimulus ideas can significantly change the outcome of an idea generation phase. Several studies highlight that the decomposition of the evaluation task by providing multi-attributive rating scales for the user also increases the accuracy of decisions [8, 9, 13, 34, 46, 47]. Moreover, Klein and Garcia [28] suggest that crowd evaluation is very helpful in detecting bad ideas, but less so when it comes to distinguishing medium or good ideas from really excellent proposals. Research also finds that facilitators of OI processes need to consider an appropriate level for users' cognitive load [9, 19]. Particularly looking at idea evaluation, there seems to occur a trade-off between accuracy and the effort users have to put into idea evaluation [60, 36]. Moreover, many studies stress the importance of the provision of rewards, incentives and other motivating elements for users [16, 34, 49, 54, 66]. For instance, users might already be inclined to participate because they can gain access to the knowledge of experts and peers [34].

Based on the results of our study, we propose a framework that reflects the current state of research (see Figure 1). The framework is based on our concept matrix (see Table 3) and includes all columns except for the study methodology, which is common across the IS discipline. The framework provides readers with a model that describes a typical OI process flow, allowing researchers to locate prior studies and structure future work more easily. It illustrates that both the sphere for the application as well as the sources of information provide the basis for an OI system. The sphere describes whether an academic, public or corporate agent is the facilitator of the OI process. This facilitator decides which source to address. Sources can either be coming from an internal crowd (such as employees) or external crowd (e.g., customers) or experts, like innovation managers or board members. The OI system itself represents an IT artifact. Researchers have to decide whether they seek to investigate full OI systems or only some parts (i.e., models, principles or methods [20, 43]). Facilitators engage their users in idea generation and/or evaluation processes. In many cases, these processes are interactive and involve user collaboration. Research investigating OI analyzes the above mentioned processes by assessing the quality, efficiency or user motivation. In what follows, we will use this framework and our concept matrix as the basis to discuss prior and identify promising areas for future research.

5 Discussion and Future Research

The finding that research on OI has most recently gained new traction underlines the timely importance of our research. Considering the vast amount of studies from various backgrounds – including many case studies –OI can arguably be considered as an important and well established means to create business innovations. In terms of idea evaluation, our study points out that researchers mostly measured the accuracy of user ratings in comparison to the evaluation of an expert committee. Though very practical, this method is also highly subjective as it depends on the expert selection and might be biased due their predispositions (e.g., having managers of a company evaluate suggestions for improvement by employees [28]). This makes the reproduction of research very difficult. Despite this disadvantage, it is a fairly common method and very suitable as many studies were case-specific and, thus, might depend on inside-knowledge from selected experts to better grasp the value of proposals.

Moreover, our literature review includes only one study that focused on the efficiency of the idea generation process [29]. However, as many firms use OI engagements, it is their employees who use the platforms for ideation. Thus, managers need to be aware of an efficient process structure in order to save valuable resources. Accordingly, one area for future research could be the efficiency of processes in idea generation. For instance, the researched we reviewed stressed that proposals are often redundant [9, 28, 46-47]. Thus, finding methods to limit similarity of ideas – for instance, through issue-based information systems – might be an interesting starting point.

Furthermore, we did not find any study that evaluated whether an internal crowd might be more accurate and efficient in delivering innovative solutions than an external crowd and vice versa. This might be another avenue for future research.

We find a number of studies analyzing idea evaluation process efficiency [e.g., 46, 8, 28, 14, 5]. However, we notice that studies on rating scales and voting techniques often decided to isolate effects triggered by social influence. For instance, both Riedl et al. [47] as well as Klein and Garcia [28] asked participants to evaluate ideas in settings where they were unable to see previous ratings by other users in order to avoid information cascades. However, in practice, users' decisions could be swayed by peer opinions [46-47, 67]. This is intuitive when looking at information sharing in social networks and, even more so, in idea markets, where facilitators explicitly build on the users' collaborative exchange of evaluations (i.e., trading activity) to derive the best ideas [32, 51, 38]. Thus, future research could investigate the robustness of different rating scales against information cascades and related effects in order to reflect more realistic conditions of OI systems.

While many studies evaluated users' motivation to participate in an OI contest, analyzing motivation was often more of a by-product rather than the main focus of any study. However, as OI becomes more professionalized, on one hand, and more of a standing, long-term process, on the other, Gassmann [17] note that motivating users becomes more challenging. Thus, future research could focus on this area as well. For instance, some studies were conducted using gamified systems, which builds on rewards, badges and other attributions to drive user motivation [49, 66, 16, 54]. However, as gamification does not necessarily lead to long-term motivation [16], future research could focus on longitude studies.

Finally, our literature review finds that extant research produced numerous models and systems for OI. However, they have rarely been adopted by other researchers. This might be related to the highly specific context to which OI processes are used for, making it difficult to generalize models and associated findings. On the other hand, idea generation, collaboration and evaluation represents a common theme in Information Systems research. There are also a number of professional OI platform providers (e.g., Hyve, Exago). Although we acknowledge the holistic approach undertaken by many studies developing a complete OI system from the ground up, we encourage future research to focus on more specific areas by contributing to the cumulative building of design theories. Gregor and Jones [20] criticized the constant re-invention of artifacts and methods under new labels, which we see happening in the literature of OI as well. The concept matrix and framework of our literature review can help to guide these approaches by providing a unified, structured approach.

This study needs to be considered against its limitations. We set strict research boundaries, following Weber and Watson [60]. Yet, this led to the exclusion of some studies from our final analysis. We might have missed some studies because they did not include the specific keywords in their title or meta-data and were not referenced by the studies we analyzed. For instance, idea evaluation can be framed as a group decision, which is a large area of IS research but is not necessarily conducted within an OI context. Furthermore, we found only few studies framing OI in the public sphere. However, as modern governments begin to involve their citizens more often in processes such as participatory budgeting [39], future research could investigate how such engagements resemble OI.

6 Conclusion

In summary, this study developed a model for research on idea generation, collaboration and evaluation in OI processes by conducting a structured literature review. We demonstrated that OI remains an emerging interdisciplinary research field, which is gaining new attention in the scientific community. Our analysis suggested that the majority of prior research investigated OI by means of case studies, often proposing an IT artifact. Our study contributes to the Information Systems literature by providing a unified, structured framework that can help to reflect and classify past research and guide future studies on OI. We also contribute to the IS literature by identifying several research gaps, which could build the basis for future research. This includes comparisons between internal and external crowds, a call for the investigation of phenomena such as information cascades, and our critique of a very limited cumulative knowledge building.

Considering the recent changes in the OI environment (e.g., accessibility for SMEs, employee-driven innovation, and professional OI platform providers), OI will most likely remain a rapidly emerging field for research. Our literature review also includes some implications for practitioners, guiding the design of future OI systems. For instance, we highlight the well-proven efficiency of multi-attributive rating scales, the acknowledgement of the users' cognitive load and the emphasis on rewards, incentives and other motivating components.

Going forward, it will be interesting to see, which mechanisms will yield the most creative and valuable ideas while still ensuring appropriate levels of effectiveness and user motivation in the long-run.

7 Acknowledgement

This study was part of the joint research project "Participation as a Service" (PaaS), funded by the German Federal Ministry of Education and Research.

References

1. ABS: Academic Journal Guide 2015, (2015).
2. Adamczyk, S., Bullinger, A.C., Möslein, K.M.: Innovation Contests: A Review, Classification and Outlook. Creat. Innov. Manag. 21, 335–360 (2012)
3. Australian Council of Professors and Heads of Information Systems: ACPHIS: Recommended IS Conferences, http://www.acphis.org.au/index.php/is-conference-ranking.
4. Bailey, B.P., Horvitz, E.: What's your idea? A case study of a grassroots innovation pipeline within a large software company. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2065–2074. , Atlanta (2010).
5. Bao, J., Sakamoto, Y., Nickerson, J. V.: Evaluating Design Solutions Using Crowds. In: Proceedings of the Americas Conference on Information Systems (2011).
6. Bayus, B.L.: Crowdsourcing New Product Ideas over Time: An Analysis of the Dell IdeaStorm Community. Manage. Sci. 59, 226–244 (2013).
7. Björk, J., Magnusson, M.: Where do good innovation ideas come from? Exploring the influence of network connectivity on innovation idea quality. J. Prod. Innov. Manag. 26, 662–670 (2009).
8. Blohm et al.: Does Collaboration among Participants Le ad to Better Ideas in IT-based Idea Competitions? An Empirical Investigation. In: Proceedings of the Annual Hawaii International Conference on System Sciences (2010).
9. Blohm, I., Riedl, C., Leimeister, J.M., Krcmar, H.: Idea Evaluation Mechanisms for Collective Intelligence in Open Innovation Communities: Do Traders Outperform Raters? In: Proceedings of the International Conference on Information Systems (ICIS) (2011).
10. Boudreau, K.J., Lacetera, N., Lakhani, K.R.: Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis. Manage. Sci. 57, 843–863 (2011).
11. Bullinger, A.C., Neyer, A.-K., Rass, M., Moeslein, K.M.: Community-based innovation contests: Where competition meets cooperation. Creat. Innov. Manag. 19, 290–303 (2010).
12. Chesbrough, H.W.: Open Innovation: The New Imperative for Creating and Profiting from Technology. (2003).
13. Ciriello, R.F., Richter, A., Schwabe, G.: Designing an Idea Screening Framework for Employee-driven Innovation. In: Proceedings of the Annual Hawaii International Conference on System Sciences. pp. 4262–4271 (2016).

14. Dean, D.L., Hender, J.M.: Identifying Quality, Novel, and Creative Ideas: Constructs and Scales for Idea Evaluation. J. Assoc. Inf. Syst. (2006).
15. Di Gangi, W.: Steal my idea! Organizational adoption of user innovations from a user innovation community: A case study of Dell IdeaStorm. Decis. Support Syst. (2009).
16. Feldmann, N., Adam, M., Bauer, M.: Using serious games for idea assessment in service innovation. In: Proceedings of the Europ. Conf. on Inf. Systems (ECIS), Tel Aviv (2014).
17. Gassmann, O., Enkel, E., Chesbrough, H.: The future of open innovation. R D Manag. 40, 213–221 (2010).
18. Girotra, K., Terwiesch, C., Ulrich, K.T.: Idea Generation and the Quality of the Best Idea. Manage. Sci. 56, 591–605 (2010).
19. Görs, J., Horton, G., Kempe, N.: A Collaborative Algorithm for Computer-Supported Idea Selection in the Front End of Innovation. In: Proceedings of the Annual Hawaii International Conference on System Sciences. pp. 217–226 (2012).
20. Gregor, S., Jones, D.: The anatomy of a design theory. J. Assoc. Inf. Syst. 8, 312–335 (2007).
21. Gressgård, L.J., Amundsen, O., Merethe Aasen, T., Hansen, K.: Use of information and communication technology to support employee-driven innovation in organizations: A knowledge management perspective. J. Knowl. Manag. 18, 633–650 (2014).
22. Horton, G., Goers, J., Knoll, S.W.: How Not to Select Ideas for Innovations: A Critique of the Scoring Method. In: Proceedings of the Annual Hawaii International Conference on System Sciences. pp. 237–246 (2016).
23. Howe, J.: The Rise of Crowdsourcing. Wired Mag. 14, 1–5 (2006).
24. Hrastinski, S., Kviselius, N.Z., Ozan, H., Edenius, M.: A review of technologies for open innovation: Characteristics and future trends. In: Proceedings of the Annual Hawaii International Conference on System Sciences (2010).
25. Hutter, K., Hautz, J., Füller, J., Mueller, J., Matzler, K.: Communitition: The tension between competition and collaboration in community-based design contests. Creat. Innov. Manag. 20, 3–21 (2011).
26. Jung, J.H., Schneider, C., Valacich, J.: Enhancing the Motivational Affordance of Information Systems: The Effects of Real-Time Performance Feedback and Goal Setting in Group Collaboration Environments. Manage. Sci. 56, 724–742 (2010).
27. Kathan, W., Hutter, K., Füller, J., Hautz, J.: Reciprocity vs. Free-Riding in Innovation Contest Communities. Creat. Innov. Manag. 24, 537–549 (2015).
28. Klein, M., Garcia, A.C.B.: High-speed idea filtering with the bag of lemons. Decis. Support Syst. 78, 39–50 (2015).
29. Kornish, L.J., Ulrich, K.T.: Opportunity Spaces in Innovation: Empirical Analysis of Large Samples of Ideas. Manage. Sci. (2011).
30. Kristensson, P., Gustafsson, A., Archer, T.: Harnessing the Creative Potential among Users. J. Prod. Innov. Manag. (2004).
31. Lakhani, K.R., Jeppesen, L.B.: Getting unusual suspects to solve R&D puzzles. Harv. Bus. Rev. 85, 30–32 (2007).
32. Lauto, G., Valentin, F.: How preference markets assist new product idea screening. Ind. Manag. Data Syst. 116, 603–619 (2016).
33. Lee, H., Seo, S.: What Determines an Agreeable and Adoptable Idea? : A Study of User Ideas on MyStarbucksIdea.com. In: Proceedings of the Annual Hawaii International Conference on System Sciences. pp. 3207–3217 (2013).
34. Leimeister, J.M., Huber, M., Bretschneider, U., Krcmar, H.: Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition. J. Manag. Inf. Syst. 26, 197–224 (2009).
35. Luo, L., Toubia, O.: Improving Online Idea Generation Platforms and Customizing the

Task Structure on the Basis of Consumers' Domain-Specific Knowledge. J. Mark. 79, 100–114 (2015).
36. Magnusson, P.R., Netz, J., Wästlund, E., W??stlund, E.: Exploring holistic intuitive idea screening in the light of formal criteria. Technovation. 34, 315–326 (2014).
37. Muller, M., Geyer, W., Soule, T., Daniels, S., Cheng, L.-T.: Crowdfunding inside the Enterprise: Employee-Initiatives for Innovation and Collaboration. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 503–512 (2013).
38. Natalicchio, A., Messeni Petruzzelli, A., Garavelli, A.C.: A literature review on markets for ideas: Emerging characteristics and unanswered questions. Technovation. 34, 65–76 (2014).
39. Niemeyer, C., Wagenknecht, T., Teubner, T., Weinhardt, C.: Participatory Crowdfunding: An Approach towards Engaging Employees and Citizens in Institutional Budgeting Decisions. In: Proceedings of the Annual Hawaii International Conference on System Sciences. pp. 2800–2808 (2016).
40. Palvia, P., Leary, D., Mao, E., Midha, V., Pinjani, P., Salam, A.F.: Research Methodologies in MIS: An Update. Commun. Assoc. Inf. Syst. 14, 526–542 (2004).
41. Pashkina, E., Indulska, M.: Where are the Ideas?: External Idea Acquisition. In: Proceedings of the Australasian Conferences on Information Systems. p. Paper 85 (2011).
42. Pedersen, J., Kocsis, D., Tripathi, A., Tarrell, A., Weerakoon, A., Tahmasbi, N., Xiong, J., Deng, W., Oh, O., De Vreede, G.J.: Conceptual foundations of crowdsourcing: A review of IS research. In: Proceedings of the Annual Hawaii International Conference on System Sciences (2013).
43. Peffers, K., Tuunanen, T., Rothenberger, M., Chatterjee, S.: A design science research methodology for information systems research. J. Manag. Inf. Syst. 24, 45–77 (2007).
44. Piller, F.T., Walcher, D.: Toolkits for idea competitions: A novel method to integrate users in new product development. R D Manag. 36, 307–318 (2006).
45. Poetz, M.K., Schreier, M.: The value of crowdsourcing: Can users really compete with profession als in generating new product ideas? J. Prod. Innov. Manag. 29, 245–256 (2012).
46. Riedl, C., Blohm, I., Leimeister, J.M., Krcmar, H.: Rating Scales for Collective Intelligence in Innovation Communities: Why Quick and Easy Decision Making Does Not Get it Right. In: Proceedings of the 31st International Conference on Information Systems, St. Louis (2010).
47. Riedl, C., Blohm, I., Leimeister, J.M., Krcmar, H.: The Effect of Rating Scales on Decision Quality and User Attitudes in Online Innovation Communities. Int. J. Electron. Commer. 17, 7–36 (2013).
48. Sawhney, M., Verona, G., Prandelli, E.: Collaborating to create: The Internet as a Platform for Customer Engagement in Product Innovation. J. Interact. Mark. (2005).
49. Scheiner, C.W.: The motivational fabric of gamified idea competitions: The evaluation of game mechanics from a longitudinal perspective. Creat. Innov. Manag. 24, 341–352 (2015).
50. Siemon, D., Rarog, T., Robra-Bissantz, S.: Semi-Automated Questions as a Cognitive Stimulus in Idea Generation. In: Proceedings of the Annual Hawaii International Conference on System Sciences. pp. 257–266 (2016).
51. Soukhoroukova, A., Spann, M., Skiera, B.: Sourcing, filtering, and evaluating new product ideas: An empirical exploration of the performance of idea markets. J. Prod. Innov. Manag. 29, 100–112 (2012).
52. Stieglitz, S., Hassannia, S.: Idea Generation by Employees and External Participants in Innovation Competitions. In: Proc. of HICCS, pp. 4272–4281 (2016).
53. Terwiesch, C., Xu, Y.: Innovation Contests, Open Innovation, and Multiagent Problem

Solving. Manage. Sci. 54, 1529–1543 (2008).
54. Toubia, O.: Idea Generation, Creativity, and Incentives. Mark. Sci. 25, 411–425 (2006).
55. Toubia, O., Florès, L.: Adaptive Idea Screening Using Consumers. Mark. Sci. 26, 342–360 (2007).
56. van Aken, J.E.: Management Research Based on the Paradigm of the Design Sciences: The Quest for Field-Tested and Grounded Technological Rules. J. Manag. Stud. 41, 219–246 (2004).
57. Wagenknecht, T., Filpe, R., Weinhardt, C.: Towards a Research Framework of Computer-Supported Organizational Participation. Springer Lect. Notes Comput. Sci. Electron. Particip. (2016).
58. Walter, T.P., Back, A.: Towards Measuring Crowdsourcing Success: An Empirical Study on Effects of External Factors in Online Idea Contests. In: MCIS 2011 Proceedings (2011)
59. Walter, T.P., Back, A.: A text mining approach to evaluate submissions to crowdsourcing contests. In: Proceedings of the Ann. Hawaii Intl. Conference on System Sciences (2013).
60. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Q. 26, xiii–xxiii (2002).
61. West, J., Bogers, M.: Leveraging External Sources of Innovation: A Review of Research on Open Innovation. J. Prod. Innov. Manag. 31, 814–831 (2013).
62. Wu, S.-C., Fang, W.: The effect of consumer-to-consumer interactions on idea generation in virtual brand community relationships. Technovation. 30, 570–581 (2010).
63. Xu, A., Bailey, B.P.: A Reference-Based Scoring Model for Increasing the Findability of Promising Ideas in Innovation Pipelines. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work. pp. 1183–1186. , Seattle (2012).
64. Yu, L.L., Nickerson, J. V.: Generating creative ideas through crowds: An experimental study of combination. In: Proceedings of the International Conference on Information Systems (ICIS). p. Paper 21 (2011).
65. Yücesan, E.: An efficient ranking and selection approach to boost the effectiveness of innovation contests. IIE Trans. 45, 751–762 (2013).
66. Zimmerling, E., Höflinger, P.J., Sandner, P., Welpe, I.M.: Increasing the Creative Output at the Fuzzy Front End of Innovation – A Concept for a Gamified Internal Enterprise Ideation Platform. In: Proceedings of the Ann. Hawaii Intl. Conf. on Sys. Sci. pp. 837–846 (2016).
67. Wagenknecht, T., Crommelinck, J., Teubner, T., Weinhardt, C.: When Life Gives You Lemons: How rating scales affect user activity and frustration in collaborative evaluation processes. In: 13th International Conference on Wirtschaftsinformatik (2017).

Systematic and Continuous Business Model Development: Design of a Repeatable Process Using the Collaboration Engineering Approach

Benedikt Simmert[1], Philipp Ebel[1], Eva Alice Christiane Bittner[2], Christoph Peters[1,3]

[1] University of Kassel, Information Systems, Kassel, Germany
{benedikt.simmert,ph.ebel,christoph.peters}@uni-kassel.de
[2] University of Hamburg, Information Systems, Hamburg, Germany
{bittner}@informatik.uni-hamburg.de
[3] University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland
{christoph.peters}@unisg.ch

Abstract. Due to permanent changes, companies constantly need to contend with new challenges. Developing and improving business models can help to adapt to constantly changing environmental conditions and to achieve competitiveness. Because most innovative developments are not the result of a single inventor, we used Collaboration Engineering to elaborate a systematic process design for business model development. To ensure an effective process design, we turned to existing knowledge by including theoretical and practical requirements of business model development. Additionally, in order to guarantee the high quality of the process, we evaluated the systematic process on the basis of a multilevel and iterative evaluation. Our evaluation clearly indicates results equivalent to expert-based business model development. Accordingly, the process design enables a continuous and recurring business model development without the ongoing support of professional facilitators.

Keywords: Business Model Development, Collaboration Engineering, Requirements of Business Model Development, Systematic Process Design

1 Introduction

Due to permanent changes, companies constantly need to contend with new challenges. Globalization and the corresponding development of the global economy bring increased transparency to the markets through new and innovative technologies. Customers have more options than ever to choose the right offer for themselves [1]. This development, in conjunction with increasingly homogenous products and services, results in an increase in the intensity of competition. Consequently, the differentiation from competitors plays an essential role for companies [2].

In this context, business models can represent an important factor to ensure competitiveness [3] and thus help to commercialize products and services [4]. Well-functioning business models can be regarded as the underlying structure for the desired

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Simmert, B.; Ebel, P.; Bittner, E.A.C.; Peters, C. (2017): Systematic and Continuous Business Model Development: Design of a Repeatable Process Using the Collaboration Engineering Approach, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 957-971

economic success of ideas, products, and services [1] and help in differentiating from competitors. Accordingly, companies can use the positive influence of business models on their performance [5].

Due to the fast changing environment of companies and the intense competition, it is necessary to develop and continually innovate the appropriate business models [3, 6]. Such a continuous refinement of its business model represents a challenging task for companies [7] and increased interest in the phenomenon [8]. In the course of the development of their business, many companies rely on established processes in the development of product innovations [9]. An empirical study from 2009 shows that business model innovation can have a greater potential for success than comparable product and process innovations [10]. With respect to these challenges, it is no longer sufficient to rely only on process and product innovation [11].

Consequently, the question arises how companies can develop business models continuously and in consistently high quality. The literature shows that the number of approaches to the development of business models has increased in recent years. Nonetheless, appropriate approaches for business model development vary greatly in relation to their procedure, their level of detail, and the techniques applied [2]. Good examples in this context are the common approaches of Osterwalder [12], Gassmann, Frankenberger, and Csik [6], Grasl [13], and Wirtz [14]. Moreover, Ebel, Bretschneider and Leimeister [15], Köster [16] and Pelzl [17] describe methods, tools, and techniques in the field of business model development. However, these approaches tend to provide higher-level concepts because of largely neglected detailed instructions and systematic process models, for example in the form of an agenda, enabling a direct implementation to the company. Against this background, the application of these approaches and tools represents a difficult task for companies, especially without the support of consultants, professional facilitators, and business model experts. Approaches that meet these criteria and provide detailed instructions and an explicit selection of methods, tools, and techniques have only been merely addressed in the existing literature. According to this reasoning, the elaboration of a systematic process design to the development of business models including detailed instructions can be seen as the next logical step in the strategic handling of business models in companies [18].

In this paper, we close the above indicated research gap and elaborate a systematic process design for business model development, which can be applied directly in companies. Because most innovations are not the result of a single person, the focus is particularly on cooperation in groups and therefore the combination of knowledge, skills, and experience [19, 20]. Against this background, we use Collaboration Engineering (CE) to elaborate a reusable collaborative group process for business model development. Collaboration Engineering offers a structured process for systematically elaborating collaborative processes [21]. With the help of this structured approach, a detailed process design for business model development will be created.

To achieve these objectives and to address the indicated research gap, we chose a tripartite research procedure. First, we deal with the question of which requirements for business model development can be identified in literature and practice. Second, we deal with the question of how to transfer the identified requirements into a systematic process design, which allows repeatability and a direct implementation. Third, we address the question of how to evaluate the effects of the application of the process design.

To answer these questions, we structured our paper as follows: We first give an overview of the existing knowledge of business model development including a working definition of the term business model. Afterwards, we explain the use of Collaboration Engineering and its applicability to business model development. To ensure an effective process design, we turned to existing knowledge by including theoretical and practical requirements of business model development derived from a literature review and an interview study. In order to elaborate the collaborative group process, we apply the central design process collaboration process design approach for Collaboration Engineering (CoPDA) [21]. In addition, we conduct a multilevel and iterative evaluation of the process design. Finally, we discuss limitations and future research and complete the paper with a conclusion.

2 Related Work

In the scientific literature, the focus has been on business models since the mid-90s. In the last 20 years, the number of scientific publications on this subject has significantly increased [5]. However, this attention is not limited to people with a scientific background. Equally, entrepreneurs, managers, investors, IT professionals, and journalists dedicated their attention to topics around business models [22]. Shafer, Smith, and Linder [23] consequently mention an anchoring of the term business model in management vocabulary.

Despite the intense debate on this issue, research in business model field is still highly active [24] and there are still different approaches, concepts, and definitions of the term business model. Accordingly, there is no universally accepted definition of the term business model [2, 5, 23, 25]. Against this backdrop, we want to introduce the definition by Wirtz [14] as a working definition.

“A business model is a simplified and aggregated representation of the relevant activities of a company. It describes how marketable information, products and/or services are generated by means of a company’s value-added component.“

The chosen definition by Wirtz is based on the value-added activities of the companies mentioned in many definitions and additionally shows which areas of value are involved in business models.

According to the rising interest on business model topics, the number of methods and approaches to the development of business models has increased [2]. Schallmo [2] notes that the level of detail in business model development approaches varies from simple descriptions to detailed process procedures. Furthermore, Rudtsch et al. [26] note that certain approaches relate to specific applications of business models (e.g., E-business models). For this reason, universal applicability is hardly feasible. Moreover, it can be stated that many approaches lack a clear connection between tools and techniques used and corresponding organizational process models for business model development [4]. This reasoning describes the need for directly implementable process designs for business model development with clear links to the tools and techniques used.

To focus on the needs of business model development and the innovativeness of groups, a detailed approach to create collaborative processes is necessary. To ensure the direct applicability of the process design, the chosen approach should provide support for the implementation of collaborative group work with appropriate tools. Collaboration Engineering meets these requirements and deals with the design and implementation of collaborative processes in order to fulfill recurring and high-value tasks. In this connection, group members combine their knowledge and skills in order to achieve a defined goal [27]. Practitioners execute the tasks of professional facilitators. In this way, there is no need for an ongoing support of professional facilitators [28].

To apply Collaboration Engineering, two important conditions, the repeatability of the task and the high quality of the task, have to be fulfilled [28]. As already mentioned, companies have to adapt to continuously changing influences. The continuous and recurring development of new business models represents an important factor for the creation of an economically successful company [3, 6]. Based on this reasoning, it can be concluded that the development of business models is a recurring task. In relation to the high quality of business model development, it is important to consider the creation of economic value through business models. In relation to the influence of business models on the performance of companies [5] and the respective strategic positioning of the business logic in the strategy of companies [29], business model development can be regarded as a high-quality task.

3 Methodology

The elaboration of the systematic process design can be divided into three sections. In order to build up the process design on existing knowledge, we first identify theoretical requirements in a literature review and practical requirements in an interview study. Based on these identified requirements of and insights into business model development, we apply the collaboration process design approach of Collaboration Engineering to elaborate the process design in a systematic manner. Additionally, in order to guarantee the high quality of the process, we evaluate the process design on the basis of a multilevel and iterative evaluation using design simulations, walk-throughs, and pilot tests. The following chapter describes the evaluation steps in detail.

3.1 Identification of Theoretical and Practical Requirements

In order to ensure an effective process design, we turned to existing knowledge by including theoretical and practical requirements of business model development. In this context, we identified theoretical requirements by means of a systematic literature review. To analyze the identified sources in a systematic manner, a category system based on the CoPDA was created. This category system reflects the CoPDA and focuses on the goals, group products, and basic conditions of the collaborative process.

As part of the systematic literature review, we looked for current journal articles that deal with the requirements of business model development in the period between 2000 and 2015. The search was conducted in the databases of EBSCOhost, SpringerLink, IEEE, and Science Direct. We used broad-based keywords (Business Model *Design, *Development, *Engineering, *Framework, *Innovation, *Process, *Tools) and eventually identified 1,256 papers. Based on a structured review process of title, keyword, and abstract search including forward and backward search, 55 relevant sources including referenced books, conference articles, and dissertations were identified. The identified conference articles had to meet the same requirements that were defined for the search of journal articles. Afterwards, the identified sources were analyzed with the help of the established category system in relation to the requirements of business model development. The corresponding results of the systematic literature review are visualized in Table 1.

Table 1: Theoretical and Practical Requirements of Business Model Development

Category	***Requirements (RQs)***	***Source***
Goals	Developing the current business model (G1)	I
	Fast and easy application of the process design (G2)	L; I
	Continuous documentation of the results (G3)	I
	Structural procedure (G4)	L; I
	Creating awareness for the need for change (G5)	L; I
Group Products	Created team spirit in the group (P1)	L; I
	Shared knowledge of basics of business model development (P2)	L; I
	Analysis of existing business model (P3)	I
	Shared knowledge about the existing business model (P4)	L; I
	Executed environmental analysis of the existing business model (P5)	L; I
	Elaborated tool/framework for business model development (P6)	I
Basic conditions	Achieve commitment (Bc1)	I
	Use a wide range of tools (post-its, index cards, mind maps) (Bc2)	L; I
	Visual representation of operating steps and results (Bc3)	L; I
	Use technical options for storing the results (Bc4)	I
	Design simple procedures (Bc5)	I
	Enable cross-divisional communication (Bc6)	I
	Convince doubters (Bc7)	I
	Arrange enough time (Bc8)	I
	Despite technology, use face-to-face approaches (Bc9)	I
	Use interdisciplinary teams (Bc10)	L; I
	Facilitators should have skills and experiences in facilitation (Bc11)	I
	Facilitators should have strong social skills (Bc12)	I
	Practitioners should have the ability for abstraction (Bc13)	L; I
	Practitioners should have strong social skills (Bc14)	I

Source: L = Literature; I = Interview study

Looking at the results in detail, it becomes apparent that the theoretical requirements identified in the literature are not sufficient to build a systematic process. First, not enough theoretical requirements were identified. Second, the theoretical requirements were not formulated with sufficient precision. To complement the theoretical requirements and to provide a more substantive basis, an interview study was carried out to identify practical requirements for business model development. Against this backdrop, eleven interviews in the context of business model development were conducted. The semi-structured interviews with experts of business model development (consultants, enterprise architects, business developers, and entrepreneurs from different industries with a minimum of three years of experience in business model development) had a duration of 30-55 minutes and were transcribed for analysis. The categories for the evaluation were deductively specified by the category system, which was also used in the literature review. By using a standardized category system, both analyses are comparable. In this context, one author defined the respective requirements with the help of an iterative and detailed coding based on a 15-step process, which was inspired by the qualitative content analysis according to Mayring [30]. Then, the results were examined and improved by the other authors with the help of a joint vote. The results of the interview study are presented in Table 1. By including theoretical and practical requirements, a detailed basis to elaborate a systematic process design for business model development is created.

3.2 Elaboration of the Process Design Using the CoPDA

In order to elaborate a systematic process design for business model development, we used the collaboration process design approach as a central design process in Collaboration Engineering. The CoPDA consists of five steps. The task diagnosis represents the first step. In this step, an analysis of required tasks, stakeholders, resources, facilitators, and practitioners is conducted. The results of this step represent defined goals and group products (outcomes) of the collaboration process. The second step addresses the task decomposition and deals with the determination of the individual activities of the process. These activities are derived from the group products of the first step. In the third step, thinkLet choice, thinkLets are assigned to each activity that has previously been identified. ThinkLets are defined as a design pattern of Collaboration Engineering. Subsequently, in the fourth step, agenda building, activities and thinkLets are transferred into an executable sequence using an agenda. The last step of the CoPDA, design validation, represents the evaluation of the developed collaborative process [21].

The elaboration of the process and the application of the CoPDA are based on the identified theoretical and practical requirements (G1-G5). In the first step, task diagnosis, the identified goals, the group products, and the basic conditions are used to define the objective of the collaborative process. The objective can subsequently be defined as follows: *"The purpose of the process design is a structured development of a business model for an observed enterprise with a heterogeneous experience and cross-functional group of up to six people in a one-day-workshop. In addition, the* *compiled results of the workshop are continuously documented. Furthermore, an awareness of the need for change is created within the group (G1-G5).”*

Based on the defined goals, group products, and basic requirements, the further steps of the CoPDA can be applied in order to elaborate the systematic process design. Thus, in the second step, the identified group products were decomposed into activities. In the third step of the CoPDA, the corresponding tasks were allocated to thinkLets, which contribute to a structured implementation of the activities. Based on this, in the fourth step, a facilitation process model (FPM) of the process was created (see Figure 2). The FPM provides an overview of the process flow. Therefore, the FPM combines the activities and the collaborative thinkLets. Thus, the elaborated process design consists of 11 activities. It should be noted that the activities of A8 to A10 are created in a loop repeating to the final drafting of the business model. To ensure the mentioned direct applicability of the process in practice, an internal agenda of the collaborative process was created in the fourth step of the CoPDA. The internal agenda shown in Table 3 offers action-guiding instructions to implement the business model development.

3.3 Evaluation of the Process Design

In order to guarantee the high quality of the process, we evaluated the systematic process design in the fifth step of the CoPDA, the so-called design validation. To perform a detailed evaluation and refinement of the process design, we chose a multilevel and iterative procedure that provides a revision of the process design according to each stage of the evaluation. Referring to this procedure, we validated the process design in four iteration loops. After each iteration loop, the process design was revised and adjusted.

To uncover hidden weaknesses, we used design simulations, walk-throughs, and pilot tests as a set of three evaluation methods. In this manner, our aim was to improve the process design continuously. Figure 1 depicts the evaluation process including the evaluation methods and the corresponding iteration loops. In the following, the evaluation is described in detail.

First, we started with a design simulation of the process design initially created. In Collaboration Engineering, design simulations represent a detailed step-by-step review of the process design by the collaboration engineer [31]. In this way, stumbling blocks in the process were identified and the formal correctness and consistency of the process were tested [21]. These improvements were directly implemented in the design simulation.

Walk-throughs are an evaluation method based on a detailed step-by-step review of the process design with experts. Walk-throughs represent the second evaluation method. In the context of walk-throughs, valuable ideas and alternative solutions are collected and discussed [32, 33]. First, we conducted two walk-throughs with experts of Collaboration Engineering. By doing so, the correct application of Collaboration Engineering was ensured and valuable suggestions for the implementation of individual activities were collected. The resulting findings were included in the process design and the second version of the process design (V2) was created. Second, we conducted two walk-throughs with business model development experts. In this way, we achieved additional insights into the facilitation of workshops on business model development. The third iteration loop was also completed by a design simulation. In this way, we created the version V3 of the process design.

As a final iteration loop and in order to check the successful application of the process design without the ongoing support of a professional facilitator, we conducted three pilot tests. The pilot tests were applied in an experimental setting in an IS master's course. In these pilot tests, the participants redeveloped an existing business model of an energy consultant platform. In this context, two of the pilot tests used the guidelines and instructions of the process design. The first pilot test was conducted by the collaboration engineer. The second pilot test was conducted by a practitioner. To compare the quality of the process design, the third pilot test was conducted by an expert of business model development. This expert did not use the guidelines and instructions of the process design. Initially, by using a questionnaire, the participants were interviewed about their previous experiences and skills in the field of business model development. Based on their experiences, the participants were randomly allocated to three groups. The pilot tests were analyzed using a questionnaire to evaluate the process design from the perspective of the practitioner. In this way "satisfaction with process", "tool difficulty", "process difficulty", "satisfaction with outcome", and the "effectiveness of the satisfaction with outcome" were examined [34–36]. The findings obtained could subsequently also be incorporated in the process. In addition, the facilitators of the pilot tests documented their experiences in a protocol. Following a final design simulation, the final version V4 of the process design was created.

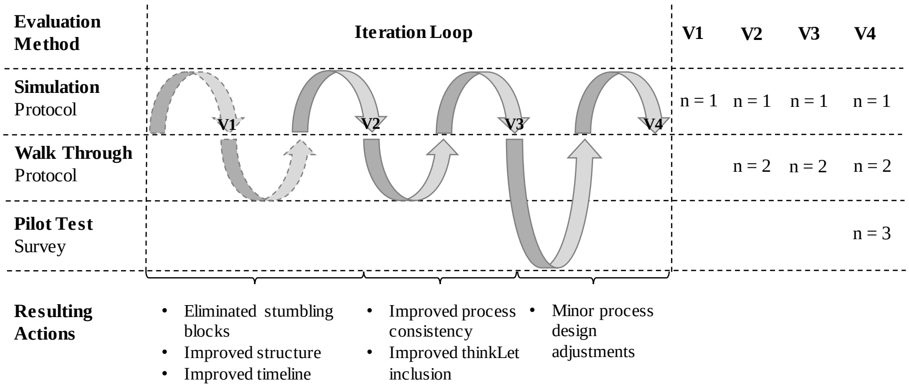


Figure 1: Iterative Evaluation of the Process Design (adapted from [37])

In the first and second iteration loop, according to V1 and V2 of the process design, mainly stumbling stones were eliminated. In the early stage, the structure of the process design was especially very fragmented. In addition, the timelines for each activity were adjusted. In this connection, CE experts highlighted the need for sufficient time for each activity. In the third iteration loop, according to V3, the focus was on the consistency of the process. In this context, thinkLets and their adoption to business model development were especially considered. Against this background, the involvement of thinkLets was increased.

The fourth iteration loop focused on the pilot test. Table 2 depicts the results and insights of the analysis of the survey. In order to fulfill the aim of CE, enabling collaborative group processes conducted by practitioners who are better or equivalent result in the comparison to professional facilitators and experts. All in all, we asked five blocks of questions. Each block consisted of five questions. In relation to the answer on a 5-point Likert scale, all groups achieved high average scores across all categories. The differences between the results are minimal and can be considered substantially equivalent. In this context, we used a simple t-test to examine if significant differences existed between the two test groups with the collaboration engineer and practitioner and the group with the business model expert. First, it should be mentioned that with exception of "satisfaction with process" of the practitioner's group and "satisfaction with outcome" of the collaboration engineer's group, the various results are not statistically significant. Considering the lack of the significance of the results (with exception of the two results mentioned), we can assume that the elaborated process design delivers results comparable to those of a professional business model expert. Thus, we can assume that the elaborated process design results in sufficiently good results, as embedded in the CE objective. In addition, the category "tool difficulty" shows especially high results. This suggests that the selection and application of the techniques used in the elaborated process design had been purposefully designed.

Table 2: Results of the Survey

Category	**Collaboration Engineer Mean (SD)**	**Practitioner Mean (SD)**	**BM Expert Mean (SD)**
Satisfaction with Process	4.51 (0,50) ^{ns}	4.31 (0,36)**	4,77 (0,21)
Tool Difficulty	4.23 (0,69) ^{ns}	4.40 (0,61) ^{ns}	4,42 (0,45)
Process Difficulty	4.25 (0,45) ^{ns}	4.06 (0,43) ^{ns}	4,11 (0,23)
Satisfaction with Outcome	3.92 (0,52)**	4.37 (0,56) ^{ns}	4,48 (0,40)
Effectiveness	4.06 (0,46) ^{ns}	4.11 (0,62) ^{ns}	4,10 (0,74)

Note: n=7 participants per group; ns = not significant, *** p<0.01, ** p<0,05, * p<0,1

Furthermore, the approximately similar results indicate that the process can also be autonomously performed by practitioners. In conclusion, we can assume that the process can be universally and directly applied in organizations without the ongoing support of a collaboration engineer or a professional facilitator. All in all, the participants of the pilot tests were satisfied with the development of the business models as well as with the results of the process.

4 Results

In this chapter, the elaborated process design is depicted and explained. As already mentioned in the part on the conceptual development, we created a facilitation process model (see Figure 2). The FPM visualizes the structured procedure with the number and name of the activity, the pattern of the collaboration, the respective thinkLets, and the suggested time of the systematic process design.

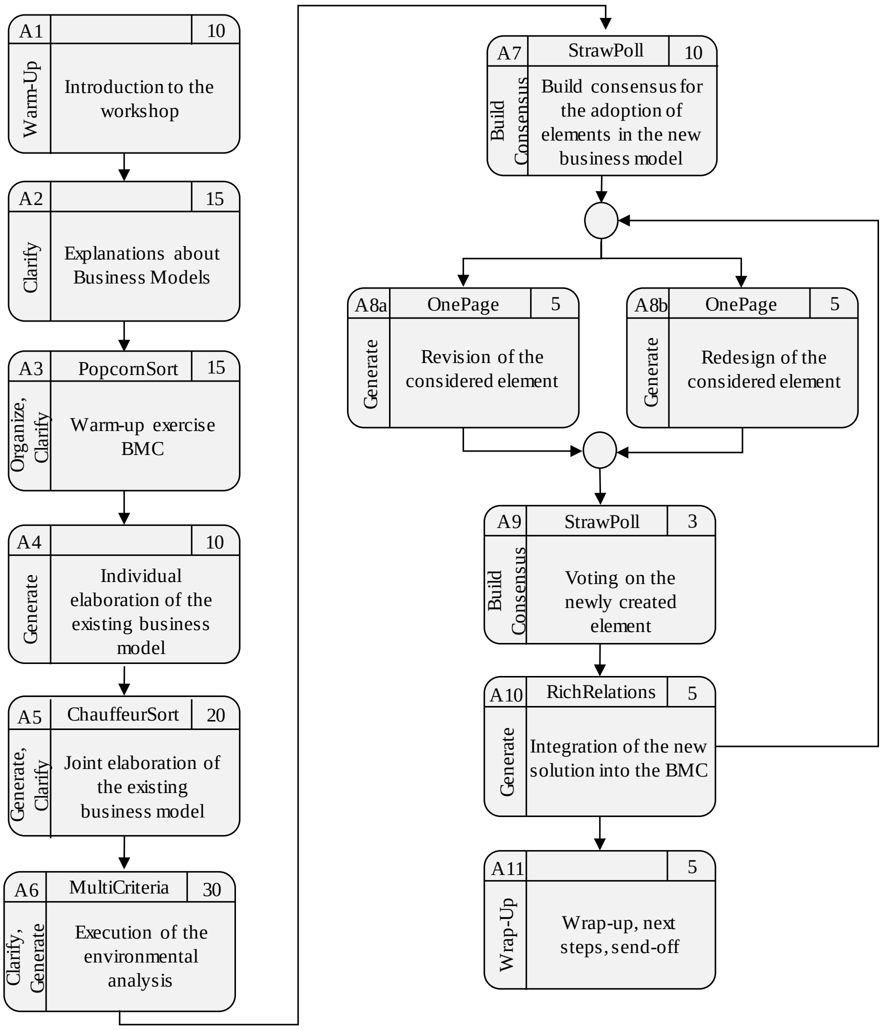


Figure 2: FPM - Systematic Process Design of Business Model Development

The internal agenda, visualized in Table 3, enables the direct implementation of the process design by providing instructions, group formations, and tools. Moreover, the internal agenda indicates how identified group products and basic conditions are incorporated into the systematic process design.

Nonetheless, not all procedures and basic conditions can be displayed in the internal agenda. In this context, the following sequence shows the respective conditions of the systematic process design.

Table 3: Internal Agenda of the Systematic Process Design

Act.	Instructions and Group Formation (PG = plenary group, SB = subgroup, I = individual)		Tools	RQs
A1	PG: Facilitator and practitioners introduce themselves. Facilitator presents the agenda and goals of the workshop. Achieve commitment to the goals from practitioners.		Presentation	P1; Bc1
A2	PG: Emphasize the relevance of BMs and explain the basic knowledge about BMs and the BMC. Ask: Do you understand the basics of business model development?		Presentation	P2; Bc1
A3	SG: Practitioners create the business model in subgroups based on content predefined in the BMC. GP: Discuss the solutions in the plenary group.		Presentation, BMC (DIN A3), prepared post-its	P2; Bc2; Bc3
A4	I: Practitioners individually elaborate the existing BM in the BMC.		Presentation, BMC (DIN A3), small post-its, pens	P3; Bc2
A5	PG: Prepare the post-its of the existing BM for the BMC. Present the post-its and discuss which field is addressed. Stick the post-its to the right place as soon as consensus has been reached. Achieve commitment and perform these steps for all predefined post-its. Summarize the existing BM and take a picture of the elaborated BMC.		Predefined post-its of existing BMC, BMC (DIN A0)	P3; P4; Bc1; Bc2; Bc3;Bc4
A6	PG: Prepare the environmental analysis questionnaire for each practitioner and introduce the practitioners to the environmental analysis (EA). I: Each participant answers the EA questionnaire (20 min.). PG: Consolidate and present the results of the EA questionnaire with the help of the EA tool. Take a picture of the results of the EA tool.		Presentation, EA questionnaire, EA tool	P5; Bc2; Bc3; Bc4
A7	PG: Facilitate the (optional) transfer of the existing elements in the new BMC. Stick the transferred elements (post-its) to a new BMC. Take a picture of the new BMC.		EA tool, BMC (DIN A0)	P5; Bc2; Bc3; Bc4
A8	PG: **Yes**: Revision of the considered element Ask: How can the considered element be revised for our business model? Orient yourself to the key questions of the respective element. The practitioners can add the existing solution and stick post-its with suggestions to the BMC.	PG: **No**: Redesign of the considered element Ask: How can the considered element be redesigned for our business model? Orient yourself to the key questions of the respective element. The practitioners are intended to stick post-its with suggestions to the BMC.	Presentation with guiding questions, BMC (DIN A0)	P6; Bc2; Bc3
A9	PG: Read each post-it of the element concerned in the BMC and ask for commitment. In case of objections, facilitate a discussion and ensure a solution (majority decision).			P6; Bc1; Bc2
A10	PG: In order to adapt the interrelations between the elements in the BMC, the facilitator gives an overview of each relationship of each element and asks for necessary additions or objections. Facilitate the discussion and ensure a solution (majority decision). Activity is performed for each element according to the order of the BMC.		Presentation with interrelation-ships, post-its, pens	P6; Bc2; Bc3
A11	PG: Summarize the workshop and the newly elaborated BM. Check if you have achieved the goals of the workshop and take a picture of the final BMC.			Bc3; Bc4

The process was consequently designed for facilitators with skills and experiences in the field of workshop facilitation (Bc11). In addition, strong social skills should be considered in practitioner and facilitator selection (Bc12; Bc14). Furthermore, the process was designed for interdisciplinary teams (Bc10). In this way, cross-divisional communication also plays an important role (Bc6). Regarding the 232 minutes of workshop time, sufficient time should be given to business model development (Bc8). Despite the ongoing technological transformation, a face-to-face approach should also be considered (Bc9). Due to the continuous obtaining of commitment, doubters could be convinced (Bc7). Altogether, the process was designed as simple as possible (Bc5).

In addition to the illustrated internal agenda, the elaborated process design contains tools and content requiring an additional explanation. Thus, the process makes use of basic knowledge of business development in general and the Business Model Canvas (BMC). In this context, it is important to prepare the basic knowledge as well as the guiding questions and the interrelationships of the BMC. Prior to the workshop, it is furthermore necessary to draw up the existing business model of the considered company. The environmental analysis plays a central role in the process design because of the consolidation of the answers to the environmental analysis questionnaire.

5 Contributions, Limitations, and Future Research

Using Collaboration Engineering and a multilevel evaluation including iteration loops, we created a recurring and directly implementable process design that contributes to business model research. As part of the elaboration of the systematic process design, we bundled theoretical and practical requirements into a systematic process design. In this way, the process design is based on the current knowledge about business model development and cooperation in collaborative processes. Consequently, we consolidated the level of knowledge in science and practice.

The innovative combination of Collaboration Engineering and business model development enables new and interesting application opportunities in the research fields of Collaboration Engineering and business model development. Moreover, this new link enables organizations to systematically develop their own existing business model by means of clearly structured instructions.

Furthermore, the individual activities in the process design represent sophisticated procedural patterns for the use and development of the Business Model Canvas. Consequently, the design process of elaborating the Business Model Canvas was transferred into a clear and structured approach.

As a result, it is possible to work up the existing business model in a structured manner and beyond represent the entire revision process using Collaboration Engineering. In summary, it is possible to adapt the business model to constantly changing environmental conditions at any time with correspondingly less preparation time. With regard to the aim of the paper, a process design that offers direct applicability and explicitly describes the use of technology has been created. Moreover, the process design enables a continuous and recurring business model development without the ongoing support of professional facilitators.

Despite these contributions, this study is not without limitations. In this context, the elaborated hypotheses of the business model are not tested in the process design. Against the background of the sample of the evaluation, additional evaluations in various contexts could confirm the effectiveness of the process design and further improve the process design itself. Accordingly, with regard to future research, the process design could be extended by a validation phase or a complementary workshop for transferring the assumptions into testable hypotheses. Another important aspect in the context of future research is the ability to create a toolbox for an individual and tailored adaptation to the needs of individual organizations. Thus, the process design can be converted into structured patterns, allowing the targeted use of individual parts of the process according to organizations' needs.

While the process is based on collaborative cooperation, the use of IT is so far underrepresented. The developed systematic process design thus serves as a basis for the further inclusion of IT in order to allow for the additional flexibility of the process design and the fast adoption of small changes. The next logical step should consider the inclusion of IT and online collaboration in the process design. The aim is to observe the underlying process design principles and to leverage the strengths of IT and online collaboration. The environmental analysis offers potential for the use of IT. In this connection, the analysis can be conducted by mobile apps or online collaboration tools to improve the process design. Moreover, additional mechanisms to combine several business models should be implemented with the help of IT.

6 Conclusion

The aim of the present paper was to create a continuous and recurring process design for business model development. In this regard, Collaboration Engineering was used to elaborate the process design. CE deals with the design and implementation of collaborative processes for the implementation of recurring and high-value tasks. The direct applicability without the ongoing support of professional facilitators characterizes the elaborated process design. In this context, theoretical and practical requirements of business model development have been identified. Based on these requirements, the systematic process design was elaborated with the help of the CoPDA. The evaluation represents another important aspect of the elaboration of the process design. Against this background, the process design was tested and improved using a multilevel and iterative evaluation. The evaluation clearly indicates results equivalent to expert-based business model development. Accordingly, the process design enables continuous and recurring business model development without the ongoing support of professional facilitators. In this context, the process design provides a detailed elaboration of the procedure steps, materials, and documents that are necessary for the facilitation and implementation. All in all, the elaborated process design created with CE can be regarded as novel way to continuously develop business models.

References

1. Teece, D.J.: Business Models, Business Strategy and Innovation. Long Range Planning 43, 172–194 (2010)

2. Schallmo, D.: Geschäftsmodell-Innovation - Grundlagen, bestehende Ansätze, methodisches Vorgehen und B2B-Geschäftsmodelle. Springer Fachmedien Wiesbaden, Wiesbaden (2013)
3. Lee, J.H., Shin, D.I., Hong, Y.S., Kim, Y.S.: Business Model Design Methodology for Innovative Product-Service Systems: A Strategic and Structured Approach. In: 2011 Annual SRII Global Conference (SRII), pp. 663–673
4. Chesbrough, H.: Business Model Innovation. Opportunities and Barriers. Long Range Planning 43, 354–363 (2010)
5. Zott, C., Amit, R., Massa, L.: The Business Model. Recent Developments and Future Research. Journal of Management 37, 1019–1042 (2011)
6. Gassmann, O., Frankenberger, K., Csik, M.: Geschäftsmodelle entwickeln. 55 innovative Konzepte mit dem St. Galler Business Model Navigator. Carl Hanser Fachbuchverlag, München (2013)
7. Palo, T., Tähtinen, J.: Networked business model development for emerging technology-based services. Industrial Marketing Management 42, 773–782 (2013)
8. Chatterjee, S.: Simple Rules for Designing Business Models. California Management Review 55, 97–124 (2013)
9. Eppler, M.J., Hoffmann, F.: Does method matter? An experiment on collaborative business model idea generation in teams. Innovation: Management, Policy & Practice 14, 388–403 (2012)
10. Lindgardt, Z., Reeves, Martin, Stalk, George, Deimler, Michael S.: Business Model Innovation: When the Game Gets Tough, Change the Game
11. Chesbrough, H.: Business model innovation. It's not just about technology anymore. Strategy & Leadership 35, 12–17 (2007)
12. Osterwalder, A., Pigneur, Y.: Business model generation. A handbook for visionaries, game changers, and challengers. Wiley, Hoboken, NJ (2010)
13. Grasl, O.: Professional Service Firms: Business Model Analysis - Method and Case Studies. Dissertation (2009)
14. Wirtz, B.W.: Business model management. Design - instruments - success factors. Gabler, Wiesbaden (2011)
15. Ebel, P., Bretschneider, U., Leimeister, J.M.: Leveraging Virtual Business Model Innovation: A Framework for Designing Busieness Model Development Tools. Informations Systems Journal (ISJ) (2016)
16. Pelzl, N.: Methodische Entwicklung von zukunftsorientierten Geschäftsmodellen im Cloud-Computing (2016)
17. Köster, O.: Systematik zur Entwicklung von Geschäftsmodellen in der Produktentstehung. Dissertation (2014)
18. Osterwalder, A., Pigneur, Y.: Designing Business Models and Similar Strategic Objects: The Contribution of IS. Journal of AIS 14, 237–244 (2013)
19. Franke, N., Shah, S.: How communities support innovative activities: an exploration of assistance and sharing among end-users. Research Policy, 157–178 (2003)
20. Sawhney, M., Verona, G., Prandelli, E.: Collaborating to create. The Internet as a platform for customer engagement in product innovation. Journal of Interactive Marketing 19, 4–17 (2005)

21. Kolfschoten, G.L., Vreede, G.-J. de: A Design Approach for Collaboration Processes. A Multimethod Design Science Study in Collaboration Engineering. Journal of Management Information Systems 26, 225–256 (2009)
22. Fielt, E.: Business Service Management - Understanding business models Acknowledgment (2011)
23. Shafer, S.M., Smith, H.J., Linder, J.C.: The power of business models. Business Horizons 48, 199–207 (2005)
24. Peters, C., Blohm, I., Leimeister, J.M.: Anatomy of Successful Business Models for Complex Services. Insights from the Telemedicine Field. Journal of Management Information Systems 32, 75–104 (2015)
25. Morris, M., Schindehutte, M., Allen, J.: The entrepreneur's business model. Toward a unified perspective. Journal of Business Research 58, 726–735 (2005)
26. Rudtsch, V., Gausemeier, J., Gesing, J., Mittag, T., Peter, S.: Pattern-based Business Model Development for Cyber-Physical Production Systems. Procedia CIRP 25, 313–319 (2014)
27. Kolfschoten, G.L., Briggs, R.O., Vreede, G.-J. de, Jacobs, P.H., Appelman, J.H.: A conceptual foundation of the thinkLet concept for Collaboration Engineering. International Journal of Human-Computer Studies 64, 611–621 (2006)
28. Kolfschoten, G., Briggs, R., Vreede, G.-J. de: Definitions in collaboration engineering. In: Monograph of the HICSS-39 Symposium on Caqse and Field Studies of Collaboration
29. Fritscher, B., Pigneur, Y.: Supporting Business Model Modelling: A Compromise between Creativity and Constraints. In: Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J.M., Mattern, F., Mitchell, J.C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B. et al. (eds.) Task Models and Diagrams for User Interface Design, 5963, pp. 28–43. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
30. Mayring, P.: Qualitative Inhaltsanalyse: Grundlagen und Techniken. Beltz Verlag, Weinheim, Basel (2010)
31. Leimeister, J.M.: Collaboration Engineering. IT-gestützte Zusammenarbeitsprozesse systematisch entwickeln und durchführen. Springer Berlin Heidelberg, Berlin, Heidelberg (2014)
32. Böhler, H.: Marktforschung. Stuttgart: Kohlhammer (2004)
33. Berekoven, L., Eckert, W., Ellenrieder, P.: Marktforschung: Methodische Grundlagen und praktische Anwendung. Gabler Verlag, Wiesbaden (2006)
34. Briggs, R.O., Reinig, B.A., Vreede, G.-J. de: Meeting Satisfaction for Technology-Supported Groups. An Empirical Validation of a Goal-Attainment Model. Small Group Research 37, 585–611 (2006)
35. Briggs, R.O., Kolfschoten, G.L., Vreede, G.-J. de, Lukosch, S., Albrecht, C.C.: Facilitator-in-a-Box. Process Support Applications to Help Practitioners Realize the Potential of Collaboration Technology. Journal of Management Information Systems 29, 159–194 (2013)
36. Kolfschoten, G.L.: Theoretical Foundations for Collaboration Engineering (2007)
37. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R.: Action Design Research. MIS Quarterly 35, 37–56 (2011)

Business Model Management: Current Practices, Required Activities and IT Support

Nicola Terrenghi[1,2], Johannes Schwarz[1,2], Christine Legner[1], Uli Eisert[2]

[1] University of Lausanne, Faculty of Business and Economics (HEC), Lausanne, Switzerland
{nicola.terrenghi,johannes.schwarz,christine.legner}@unil.ch
[2] SAP AG, Innovation Center Network, St.Gallen, Switzerland
{nicola.terrenghi,johannes.schwarz,uli.eisert}@sap.com

Abstract. This paper explores the concept of business model management, defined as a generic process covering all phases of the business model life-cycle. In contrast to previous business model literature, which is mainly focused on the design of business models, we argue that the successful exploitation of the business model concept requires a dedicated management approach. Due to the lack of extant research in the domain, we build on multiple, exploratory case studies of large organizations, based on 20 expert interviews. This paper contributes to a better understanding of the current practices and needs in business model management and the multifaceted role of the business model concept in each of its phases. Moreover, we suggest roles of IT in the business model management process.

Keywords: Business model management, business model design, business model implementation, case study

1 Introduction

A business model (BM) is an abstract representation of business logic [1]. Serving as a reference framework, it supports practitioners in conceiving, designing and communicating business ideas [2–4]. The academic literature provides analyses of how organizations design and innovate their BMs [5–7]. However, although the BM can be understood as a structured management tool [8], there is still no clear understanding of its roles beyond design and innovation, also seen as the transition from BM *plan* to its *execution* [9].

"A 'strong' business model can be managed badly and fail, just as much as a 'weak' business model may succeed because of strong management and implementation skills" [9]. For instance, Ryanair's BM "creates several virtuous cycles that maximize its profits through increasingly low costs and prices" [10]. Its competitive advantage keeps growing as long as managers make these virtuous cycles spin. Being designed in a dynamic environment, as a reaction for instance to market changes, increasing competition, or technological innovation, BMs thus require active management [11]. Although recent research indicates that the BM life-cycle should embrace more phases than just design [12], these phases and a holistic management

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Terrenghi, N.; Schwarz, J.; Legner, C.; Eisert, U. (2017): Business Model Management: Current Practices, Required Activities and IT Support, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 972-986

approach have seen very little attention in academia and practice [13] – wrongly, in our view, because the BM concept can be leveraged by decision makers for implementation and management of a successful business, [9] and shifts "from a conceptual and theoretical focus to tooling and practical usability" [14]. We address this gap by answering the following research questions: (RQ1) *What is the current state of practice in business model management?* (RQ2) *What are the roles of the business model concept in each phase of this process?* If BMs are found to be relevant from a management perspective [15], decision makers also need IT-based BM tools to support their management. To address this aspect, we ask a third question: (RQ3) *What are IT's roles in supporting the business model management process?* Building on a qualitative research design and insights from 20 case studies, our findings confirm that companies mostly use the BM concept for analysis and design, but have not fully embraced it as management instrument in the implementation and control phases. Our study provides a baseline for future research by providing conceptual foundations and developing avenues for practice and future research.

The remainder of this paper is structured as follows: We start by reviewing the literature to derive an analysis framework for BM management (Section 2). We then present our qualitative research design (Section 3). Our empirical insights shed light on the current practices and required activities in 20 cases (Section 4). We conclude by developing avenues for practice and future research (Section 5).

2 Theoretical Background

In this chapter, we review the BM literature and elaborate on the emerging theme of BM management. Building on our literature review, we suggest an analysis framework that will guide our empirical study.

2.1 Business Model Concept

Although every company adopts a BM, either explicitly or implicitly, it remains an open question what exactly is understood by BM, that is, how it should be conceptualized [16]. At a fundamental level, scholars and practitioners agree that the BM is crucial for the success of today's organizations, especially concerning growth potential [17], competitive advantage [18, 19], long-term performance [20], and as a new source of innovation [21]. The academic community has recognized BM's potential, leading to a rapidly growing number of publications across management, IT, strategy, and other disciplines [8, 22].

In practice, the growing interest in the concept was catalyzed by the popularity of representations of the BM such as the so called Business Model Canvas [3], e3-value ontology [2], or St.Gallen Business Model Navigator [23], which have become reference tools for business innovation in both entrepreneurial and large organizations [14]. The main benefits of these BM conceptualizations are the systemic and abstract representations of the core business logic [9, 1, 24], which can serve as a useful device to analyze and communicate a company's value creation, delivery, and capture mechanisms [16].

2.2 Business Model Management

Recently, scholars have begun to investigate the phases of BM management beyond design, such as analysis [25], implementation [26], and management [27]. However, the literature provides very limited and inconsistent conceptual or empirical evidence on the practices and characteristics of BM management as a holistic process [17, 28]. For instance, [35] identifies design, implementation, operation, change, and controlling phases within the so-called business model life-cycle. [30] document an experimentation and exploration phase, followed by an implementation stage, and [23] distinguish between the phases of mobilizing, understanding, designing, implementing, and managing BMs.

Generally, the BM concept can be considered either as a static *blueprint* or as transformational, "using the concept as a tool to address change and innovation in the organization, or the model itself" [29]. The latter view reflects a management context, because it considers the BM as a tool to support the analysis, planning, and transformation of organizations [30]. It also comprises the transformation of the BM itself. Specifically, a business model "is typically a complex set of interdependent routines that is discovered, adjusted, and fine-tuned by 'doing' " [29]. Given how crucial it is to sustain the BM in volatile ecosystems, all tasks should be managed deliberately and its "conception, introduction into the marketplace, and ongoing management should not be left to chance" [31].

2.3 IT Support

Software and tools that support BM management have already been developed but are still in the earliest state of immaturity and are largely restricted to the design phase, by supporting the visualization of a BM [15]. However, recent research highlights that such "computer-aided business modelling tools should go beyond simple design tools and evolve into an own class of high-level decision support tools" [27]. This implies that software tools not only support design, but also the overall process.

2.4 Towards a Framework of Analysis for Business Model Management

Building on the literature, we develop an analysis framework to guide our empirical research on BM management (Figure 1). We suggest defining BM management as *an generic management process, building on the business model as the central unit of analysis.* The overarching element of our framework is the BM conceptualization employed in practice. Through this element, we account for a better understanding of the relevant aspects of the BM concept in the different BM life-cycle phases. We refer to specific BM conceptualization and operationalization as well as the general mindset and perception of practitioners toward the BM concept. For BM management, we rely on four generic phases of strategic management: analysis, design, implementation, and control [32].

Analysis. Given that the development of new BMs or the improvement of existing ones is a complex challenge for managers [33], understanding when changes in the BM are needed is essential [34]. During this phase, relevant aspects such as technological innovations in the external and internal environment are identified to assess the urgency of and opportunities to alter the BM [34].

Design. "Designing a new business model requires creativity, insight, and a good deal of customer, competitor and supplier information and intelligence" [17]. Design could include activities such as brainstorming, prototyping, testing, and selection [3] and could be theoretically guided by dedicated frameworks [24].

Implementation. Design and implementation phases are distinct but strongly related: a 'good' BM design can lack the expected value owing to its ineffective implementation [9]. Thus, in this phase, project management could become relevant [12] in order to operationalize the BM [26].

Control. Although the BM may have been rigorously designed and implemented, its de facto performance and effectiveness are subject to emerging events and needs being continuously controlled, for instance, in relation to financial, process, or growth performance [35] and associated risks [16].

These phases can be mapped to the ones mentioned in [3, 9, 12] and reflect the more generic phases of strategic management. However, these four phases cannot be considered as exact, given the early stage of BM management. They should be considered as 'idealized' rather than 'ideal'. In line with the literature, phases could be carried out in parallel, iteratively, or ad-hoc [3, 9, 12].

The underlying element is IT support (similarly to [15]), because each BM management phase can be expected to be supported by one or more software tools [15].

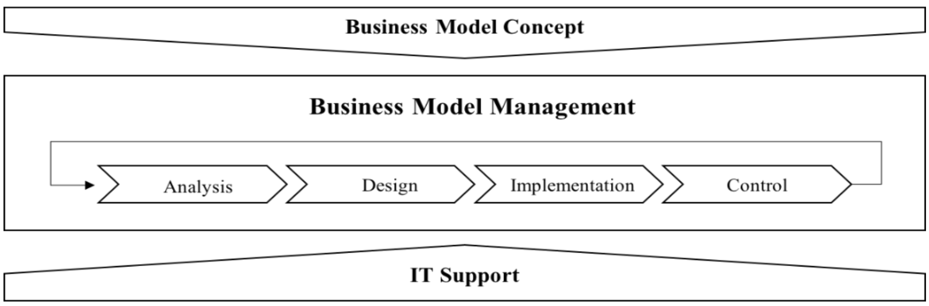


Figure 1. Framework of Analysis for Business Model Management

3 Methodology

To better understand the current situation and roles of BM management in practice, we applied a qualitative and exploratory research design by conducting multiple case studies [36]. A case study method is commonly defined as "empirical inquiries that investigate a contemporary phenomenon within its real-life context"; it is particularly useful to understanding the emerging topic of BM management [36]. We performed multiple case studies because this approach reinforces the generalizability of results [37]. We conducted a series of 20 cross-industry cases of large organizations. For each case, we collected data through semi-structured interviews with experts in different managerial roles.

3.1 Case and Expert Selection

To gain insights into BM management in organizations, we had to identify suitable case sites, as well as experts as key informants. Our goal was to cover a broad variety of organizations from different industries and with different levels of expertise in BM management in our case sample.

Table 1. Characteristics of Interviewees and Their Organizations

Roles	Managers (18), Director (1), CEO (1)
Divisions	Innovation (12), corporate strategy (6), business development (2)
Industries (code)	Automotive (AUT) (2), financial services (FIN) (2), energy (ENE) (3), chemicals (CHE) (3), conglomerate (CON) (3), information technology (INF) (2), logistics (LOG) (1), high-technology (HIG) (3), research (RES) (1)
Countries	Germany (11), Switzerland (3), France (2), U.S. (2), Italy (1), Norway (1)
Revenues (avg.)	$40B
Size (avg.)	98K employees

Before approaching a wider number of companies, we conducted four knowledge-sharing sessions with affiliated practitioners who have experience in BMs in their organizations and were willing to share their insights. Based on the results of the knowledge-sharing sessions, we learnt that the topic is particularly relevant for large organizations (> 250 employees) across industries: compared to small- medium-enterprises, they tend to assign more resources to BM-related initiatives (e.g,. dedicated teams or organizational units). Further, decision makers in large organizations are usually aware of the BM concept, but struggle to manage it. The knowledge-sharing sessions have been also crucial to spot roles of 'BM experts' in an organization: we identified experts in several areas such as new business development, innovation management, general management, and strategic management. Finally, we approached potential candidates at conferences, events and through personal contacts of the authors and seven colleagues. We invited decision makers in large organizations who have BM knowledge and have adopted or intend to adopt it in the management of one or more phases. Of the 34 experts we invited, 24 agreed to be interviewed; all had managerial or executive roles in large organizations. Four could not provide any relevant insight on BM management phases and were excluded, resulting in 20 eligible cases. Table 1 summarizes the characteristics of the organizations and interviewees. To guarantee anonymity, we refer to each expert with a code (e.g., AUT1 = the industry of the organization s/he belongs to: Automotive 1).

3.2 Data Collection and Analysis

Between January and May 2016, we conducted 20 semi-structured interviews by the authors or one of their seven colleagues. The interviews were done face-to-face (18 of 20) or via telephone (2 of 20) and lasted between one and four hours. In some cases, a first brief telephone interview was followed by a meet-up. Usually, between one and three interviewees participated, often from very different organizational units, reinforcing the topic's relevance across divisions. At least two interviewers participated; one leading the conversation and the other(s) taking notes. Owing to confidentiality, we did not audio-record any session; statements in quotation marks can therefore not be considered as direct citations. During the interview process, we presented our current understanding of BM management to interviewees, followed by questions about their previous experience with and use of BMs, the topic's general relevance, IT's potential roles, and general feedback, followed by an open discussion.

For data analysis, the interview notes (between pp. 2 to 5 per interview) were consolidated and analyzed by two authors, individually, along the process of constant comparison and iterative conceptualization [38]. Some interviews were done in German or Italian and had to be translated into English. In the analysis process, we categorized the codes into the elements of the theoretical framework before identifying emerging themes in each group. In case of disagreement, a third author was involved in the discussion until we had clarification. Finally, each author reviewed the coding, and we achieved agreement in a round table session.

4 Results

4.1 The Business Model Concept

Regarding the BM concept, we highlight two important findings: its operationalization and its relevance across BM management phases. While some academics argue about the 'best' BM conceptualizations, our results suggest that certain organizations are able to leverage the BM concept without specific templates (AUT1, FIN1, FIN2, CHE3), while others operationalize the concept through existing BM frameworks, such as Osterwalder's Business Model Canvas (CON2, INF2, CHE2). However, there appears to be consensus that, sooner or later, organizations adapt these frameworks to their specific needs and/or integrate them with other templates (e.g., SWOT). As CHE1 notes, "*we need flexible canvases and tools, because each company develops and uses their own.*" CON3 confirms that "*every company has its own specifications, which must be represented.*"

Regardless of its specific conceptualization, the relevance of the BM concept differs significantly across the BM management phases. Organizations strongly apply a BM perspective during the design phase and highlight its importance, such as "*visual representations*" (CHE2), "*spot and rank the most critical elements of a BM through the eyes of the customers*" (AUT2) or "*represent alternative BM designs*" (CON3). However, they report additional issues during implementation. Some companies report that "*other divisions do not speak this language*" (AUT2), "*some colleagues don't want to think creatively*" (CHE1), and "*management looks at ROI*" (INF2). Since more stakeholders are increasingly affected in advancement throughout the phases (RES1, CHE3), conflicts with the core business must be avoided (AUT1).

4.2 Business Model Management

Table 2. Results Per and Across BM Management Phases (insight recurrence)

	Analysis	*Design*	*Implementation*	*Control*
Current practices	Environment analysis (7)	Iterative testing (6)	Accountability management (3)	Financial performance (2)
	Partnerships (3)	Customer centricity (4)		
Required activities	Prevention of disruption (5)	Stimulation of scenarios (6)	Stakeholder management (4)	Holistic dashboard (6)
	Portfolio analysis (3)	Enhanced visualization (4)	BM portfolio management (2)	Prompt notification (4)
	Idea management (2)			
IT support	Key value to current practices and required activities (6) Creativity and communications are enabler, not inhibitors (4) Integrated with other tools (3)			

Analysis. *Current practices.* The analysis phase is mostly described as the process of collecting relevant information from the external environment, which refers not only to customers (CON1) but also competitors (AUT1, LOG1), other industries (ENE2), new market entrants that "*might intervene and disrupt the market*" (CHE1), and digital trends (CHE1, INF1, ENE2). Some organizations proactively analyze their environment, adopting an "*intensive collaboration with startups*" (ENE1) to "*help them grow and to change [their] business model based on their insights*" (CHE1). Similarly, CHE2 mentions that her organization constantly runs internal workshops and "intrapreneurship" programs to help employees generate and collect business ideas.

Required activities. The lens on the external environment is primarily based on the question "*how could my business model be disrupted?*" (CHE3). This dilemma concerns the experts' hope, and fear, that "*digital transformation might change things or open new opportunities*" (ENE2). Thus, interviewees highlight the need for a prompt signal that can spot a disruption threat, or opportunity, in advance, because "*it is […] important to find the right timing for business model innovation.*"

Among the organizations that trigger BM design based on the analysis of internal capabilities (HIG2, CHE2, CON3), the primary activities concern the "*understanding of portfolio of [existing] business models*" (CON3). The same expert highlights that "*Not every product must be a financial performer, because it could complement services. It is important to map these dependencies. We would always start on the portfolio layer.*" The collected information is often unstructured and lacks a filtering process that leads to the design phase. For this reason, CHE2 suggests that an "*idea management tool*" is needed. Thus, the crucial activities needed in the analysis phase concern the identification of possible changes in the industry, systematic management of insights for BM design or adaptation, and the acknowledgment of internal constraints relating to existing BMs.

Design. *Current practices.* BM design should use the customer as starting point for value creation. Experts tend to validate their assumptions through market analysis and product test (e.g., prototypes). For instance, CHE3 states that they "*test different elements of the [business model] canvas with customers (e.g., willingness to pay).*" In this sense, the customer is perceived as a co-innovator. AUT1 highlights the relevance and urgency of customer centricity owing to digital transformation, suggesting that design "*needs to anticipate the digital expectations of customers three years ahead.*" However, although most organizations highlight customer centricity, an expert in the oil and gas industry (ENE3) states that in their commoditized market, this is not relevant, since "*we just need to find resources and sell them.*"

A second recurrent – and expected – insight is that the BM design is typically iterative. Experts highlight the essential, repetitive validation of their assumptions, which can take place through the adoption of specific tests (CHE1, CON3) and through prototyping (CHE3, HIG1, CHE3). Tests are considered particularly valuable for generating multiple BM designs, also known as *versioning* (CHE2). Prototyping is a practice that, according to CHE3, helps to "*avoid long-term investments upfront.*" However, two experts mention feasibility regarding testing and prototyping: ENE3 and RES1 highlight that constraints such as time, legislation, and infrastructure often prevent testing and prototyping.

Required activities. Although testing and prototyping are currently common practices for several organizations, such as minimum viable propositions and surveys, interviewees mentioned the need for further support in the validation of BM design. In particular, "*business model simulation*" is a recurrent term (e.g., AUT2, INF1). CHE1 describes it as the assessment of "*internal and external requirements, so that the business model can work in particular situation[s] or market[s].*" Similarly, INF1 focuses on the relevance of simulations to predict "*the best customer channel, sales and distribution.*" Such activity is also highlighted as an IT solution to support managers in creating multiple scenarios or *versions* of a BM (CHE2) and to identify the critical elements of a BM (AUT2).

To support BM simulation and validation, experts suggest that a detailed visualization of a BM and its elements is needed: CHE2 states the "*need to get more* *details in the business model canvas to actually use it and communicate ideas.*" AUT2 also supports the urgency for a "*granular*" representation that "*avoids ping-pong in the communication between departments*," while FIN2 focuses on a BM visualization that enables better customer segmentation and description.

Implementation. *Current practices.* The implementation phase holds unique characteristics, particularly compared to the design phase. Generally, it appears to be "*a challenge at a large company*" (HIG1), perhaps even "*the biggest hurdle*" (HIG2). Specifically, the first important characteristic noted by several interviewees is the increasingly significant role of stakeholders. Employees who are directly responsible for BM implementation and execution are a primary stakeholder group. CON3 describes this group as the "*entrepreneurial team*" and asks "*what are the ideal characteristics of the members of this team?"*

Required activities. Concerning stakeholders, ENE3 addresses the "*need to have a business model owner, who pushes the execution with the implementation team.*" As several other respondents highlighted the need for general stakeholder management, we see a second stakeholders group beyond the BM owner. These stakeholders could be the implementation team (ENE3), sponsors (CHE3, HIG1), or coaches (CON3). Stakeholder management's role is also emphasized by the need for "*approval management*" (HIG1, CH1, FIN1) or "*stage-gate processes*" (RES1). ENE3 suggests employing a "*power couple*" – two managers accountable for BM design and implementation respectively, and who strongly depend on each other.

While the previous phase indicated a focus on finding the right BM and assessing customer needs, implementation is concerned with finding the right organizational set-up. Several companies mentioned potential conflicts between the traditional core business and new BMs. For instance, AUT1 noted that "*the central question is how the traditional core business can be combined with new, digital businesses*," facing major challenges such as self-cannibalization, which must be avoided, because "*there is no willingness to give up existing business for new ones*" (CON1).

Control. *Current practices.* Only two interviewees addressed the regular control of one or more implemented and 'running' BMs (INF1, FIN2). They highlight continuous monitoring of financial performance, referring to specific BM elements, such as revenues and costs. Both INF1 and FIN2 suggest that there is currently no control tool besides financial reports and, when necessary, ad hoc analysis of specific reports' insights.

Required activities. CON3 argues that "*business models are never stable.*" To control a BM, experts suggest that data and IT systems are necessary. Here, INF2 comments that "*[they] need to make a dashboard*" to visualize and leverage quantitative data for BM improvement. Similarly, CHE2 states the "*need to look at environment data*" to monitor and prevent threats and seize opportunities, while CON1 asks "*if the customer has an input, how [can we] leverage this input in the business model in the best way?*" INF1, also owing to his industry, suggests that such a dashboard should provide "*real-time monitoring*" of the BM.

The regular controlling of a BM and its visualization as a dashboard form the basis for working as "*early warning*". CON3 explains that "*alerts must notify business model owners if crucial parameters deflect. Thus, these parameters must be captured (for instance, financial indicators, regulations, or other requirements).*" FIN2 argues that his organization needs to enhance reporting practices in order to enable more complex queries. RES1 and CHE3 discuss how alerts should also help managers to distinguish between the need for incremental change and the need for disruptive change.

4.3 IT Support for Business Model Management

While most respondents reported the need for general software support of BM management, only CON3 is already systematically working with a mix of "*predefined Excel files for evaluation and risk management, as well as PowerPoint, and one tool to describe processes.*" With two exceptions (ENE2, HIG1), most companies argue – similarly to CHE1 – that "*software support would be a key additional value*" (CON1, CHE1, CHE2, RES1, CON2, FIN1). For instance, CON1 states that "*no tools are used, which is a big problem*," while CON2 notes that "*we need situation-specific and iterative tools.*" However, a major requirement of those tools is that they could be helpful by providing a pre-structured process that "*moderates the process, colleagues, and process sequence*" (FIN1) yet maintains the agility and flexibility, particularly during the design phase. "*We need […] iterative tools to avoid fixed step-by-step processes with unnecessary steps*" (CON2).

While some firms are concerned about creativity and agility during design, the results show that the implementation phase could strongly benefit from software support. The primary reasons are that implementation requires support for "*identifying experts in the network*" (CHE1) and "*identifying organizational units that could provide knowledge or support the business model to implement it faster*" (CON3). From the interviews, we also learnt that BM representation should "*provide deep dives into certain elements*" (CON3), an "*ecosystem view*" (RES1), and be "*integrated with other conceptual tools, such as strategy-maps or balanced scorecard*" (FIN1). We can also report two significant relationships with existing software systems (CON1): "*a BM perspective could be integrated into existing tools, such as CRM*" and "*how can business models be derived from the ERP?*"

5 Discussion

Our study contributes to the BM literature by suggesting a comprehensive conceptualization of BM management. Based on empirical insights from 20 cases, we assessed current practices and required activities in BM management. Our results reveal that organizations still concentrate on BM analysis and design, but also confirm the need for a structured management process. The above mentioned empirical findings outline future avenues for research and practice, laying the groundwork for further studies.

5.1 Business Model Management

In response to research question 1, our results suggest that current BM practices in large organizations mainly relate to analysis and design, and that BM management, as a holistic management approach, is highly relevant to companies across different industries. In this sense, it empirically confirms previous theoretical arguments [3, 13], which hold that the BM concept is widely accepted for designing and innovating new BMs but that there is a significant need to focus on the BM management as a whole, including implementation and control (Table 3).

Our study is also one of the first to explore the details of such a holistic process in terms of current practices, required activities, and IT's roles. The results show that organizations have an expressed need for a general process to manage the entire life-cycle of BMs and provide a first rich picture of the current situation in practice. Further, the maturity level, which could be considered as experience with BMs and the level of adoption of the concept in different phases, varies widely. Very few organizations have already adopted a fairly structured BM management process (CON3, CHE1), while most either begin with small BM-related projects or have integrated the BM perspective into their existing innovation processes (e.g., product innovation).

5.2 The Roles of the Business Model Concept

Regarding research question 2, we found that the BM concept is in fact important during all phases. Interestingly, its role and application appears to change along the phases, pointing at a multi-faceted conceptualization (Table 3). Our results show for instance that the *working mode* changes along these phases. During analysis and design, frequent iterations, agility, simulations, and customer centricity are key. During implementation and control, a fairly sequential procedure with "*approval management*" (CHE1) and "*milestones*" (CON3) was indicated. In addition, the main sources of knowledge also differ among the phases: during the analysis, the environment is crucial; in the design phase, the focus is on products and customers; the implementation phase depends strongly on the contribution of internal experts and on the integration with other internal BMs; the control phase requires data acquired from enterprise systems and other data sources. Irrespective of the phases, some authors even go beyond the management of a single or multiple BMs; they state that the BM concept is a crucial perspective to manage the entire corporation, irrespective of the company's size. Such "*evolution*" of the BM concept has been recently coined by [39] as *business model-based management*.

A primary obstacle to the further adoption of the concept, particularly during implementation and control phases, could be the "*lack of business model mindset*" (INF1, AUT1, CHE2, ENE1, ENE3, FIN2). Here, many divisions such as corporate strategy become involved. According to the experts' insights, these stakeholders may have a short-term perspective (FIN2), looking solely at financial performance and ROI (CHE3, INF2) and may feel threatened by changes to the traditional BM (AUT2).

Table 3. Overview of the Characteristics of Business Model Concept and Management

	Analysis	*Design*	*Implementation*	*Control*
Maturity level (current practices)	High (10)	High (10)	Low (3)	Low (2)
Working mode	Explorative, open	Agile, iterative	Gateways, sequential	Structured, regular
Key knowledge source	Environment	Offer, customers	Internal experts, other internal BMs	Internal data source
Obstacles to the adoption of BM management	Lack of BM mindset; focus on short-term results; fear of cannibalizing the traditional BM			

5.3 IT's Roles

Consistent support of the entire BM life-cycle raises the need for adequate IT solutions, not only as digital visualization of BMs, where pen-and-paper seems to be sufficient for most practitioners. We respond therefore to the quest for research on "*IT support for developing and managing business models*" [15] by providing first insights into requirements of the prospective research stream on software systems that "clearly go beyond simple design tools […] and evolve into an own class of high-level decision support tools" [15]. IT must support a structured management process, but should not constrain agility in and the iterative nature of the early phases of BM development. Handling this paradox could be a fundamental aspect of IT support to drive the adoption of systems and to ensure their sustained use throughout the life-cycle. In other words, paying attention to the above mentioned characteristics of the different phases may contradict the over-simplistic assumptions of BM software tools, which mainly focus on visualization. Further, a selection of specific features was highlighted, namely simulation, collaboration, knowledge management, reporting, and the fact that IT support should include additional strategy tools such as strategy-maps that are easy to adapt.

We also contribute to the broader IS literature by building on [27]'s argument about IS's key roles in "informing strategic disciplines and in contributing to increase understanding of the essence of BMs and other strategic notions." While [27] suggest three primary contributions of IS to strategic management (modeling at a strategic level, strategizing as designing, considering computer-aided design), we added an additional role of IS. Given the high relevance of implementation for any BM's success, a fruitful future research area could be to investigate to what extent the existing IT-landscape is a barrier to successful BM management. For instance, CON1 notes that not all BM ideas can be implemented, owing to the rigidity of the IT-landscape, and that it should be examined how existing IT systems (e.g., ERP, CRM) can support BM management, providing data or integrating a BM perspective.

6 Conclusions

This study investigates the current practices and required activities associated with BM management. Drawing on empirical insights from 20 case studies, we found evidence that managers in large organizations acknowledge the BM concept's relevance, not only for the purpose of design, as already established in the literature, but also for analysis, implementation, and control. We have shown the increasing relevance of the BM concept throughout all phases and, although companies report different maturity and adoption levels, it affects different activities and multiple stakeholders.

Our study provides a baseline of BM management from a practitioner perspective, and we trust that it will inspire other researchers to contribute to this emerging research stream. Specifically, we identified three primary limitations that could trigger future studies. First, we approached practitioners interested in the topic, rather than those who would deliberately not manage BMs (e.g., managers in controlling, who have all the tools they need in place). Interviewing such a group could provide useful insights into the barriers to BM management adoption and could therefore sharpen BM management's value proposition.

Second, BM management is still in its infancy. Organizations use the concept mainly for design and creativity purposes (e.g., during workshops) or to complement other innovation processes and therefore have little experience with BM implementation and control. It follows that our findings on these two phases need for further validation. In particular, corporate spinoffs or recently introduced BMs might be interesting subjects for case studies.

Finally, in this study, we sought to obtain a broad overview of the state-of-the-art in BM management. Thus, we approached a fairly large number of organizations. Future research could build on our findings to conduct in-depth research into selected organizations that adopt BM management in each phase.

References

1. Al-Debei, M.M., El-Haddadeh, R., Avison, D.: Defining the Business Model in the New World of Digital Business. Proc. Fourteenth Am. Conf. Inf. Syst. 1–11 (2008).
2. Gordijn, J., Akkermans, H.: Designing and evaluating E-business models. IEEE Intell. Syst. Their Appl. 16, 11–17 (2001).
3. Osterwalder, A., Pigneur, Y., Clark, T., Smith., A.: Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers. (2010).
4. Voigt, M., Plattfaut, R., Ortbach, K., Malsbender, A., Niehaves, B., Becker, J.: Evaluating business modeling tools from a creativity support system perspective - Results from a focus group in the software development industry. Proc. - Pacific Asia Conf. Inf. Syst. PACIS 2013. (2013).
5. Chesbrough, H., Rosenbloom, R.S.: The role of the business model in capturing value from innovation : evidence from Xerox Corporation' s technology spin-off companies. Ind. Corp. Chang. 11, 529–555 (2002).

6. Massa, L., Tucci, C.L.: Business Model Innovation. The Oxford Handbook of Management Innovation (2013).
7. Baden-Fuller, C., Haefliger, S.: Business Models and Technological Innovation. Long Range Plann. 46, 419–426 (2013).
8. Wirtz, B.W., Pistoia, A., Ullrich, S., Göttel, V.: Business Models: Origin, Development and Future Research Perspectives. Long Range Plann. 49, (2015).
9. Osterwalder, A., Pigneur, Y., Tucci, C.L.: Clarifying business models: origins, present, and future of the concept. Commun. Assoc. Inf. Syst. 15, 1–43 (2005).
10. Casadesus-Masanell, R., Ricart, J.E.: How to Design a Winning Business Model, https://hbr.org/2011/01/how-to-design-a-winning-business-model. (Accessed: 23.07.2016)
11. Bouwman, H., Faber, E., Haaker, T., Kijl, B., de Reuver, M.: Conceptualizing the STOF Model. In: Bouwman, H., De Vos, H., and Haaker, T. (eds.) Mobile Service Innovation and Business Models. pp. 31–70. Springer Berlin Heidelberg, Berlin, Heidelberg (2008).
12. Wirtz, B.W.: Business Model Management. Zeitschrift für Betriebswirtschaft. 81, 351–353 (2011).
13. Kijl, B., Boersma, D.: Developing a business model engineering & experimentation tool - The quest for scalable lollapalooza confluence patterns. AMCIS 2010 Proc. 662, 78–80 (2010).
14. Bouwman, H., Reuver, M. De, Solaimani, S., Daas, D., Haaker, T., Janssen, W., Iske, P., Walenkamp, B.: Business Models Tooling and a Research Agenda Harry. 25th BLED eConference. 1–28 (2012).
15. Veit, D., Loos, P., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Spann, M., Leimeister, J.M.: Business Models: An Information Systems Research Agenda. Bus. Inf. Syst. Eng. BISE (2014).
16. Burkhart, T., Krumeich, J., Weth, D., Loos, P.: Analyzing the Business Model Concept—A Comprhensive Classification of Literature. In: Thirty Second International Conference on Information Systems. ICIS pp. 1–19 (2011).
17. Teece, D.J.: Business models, business strategy and innovation. Long Range Plann. 43, 172–194 (2010).
18. Mitchell, D., Coles, C.: The ultimate competitive advantage of continuing business model innovation. J. Bus. Strategy. 24, 15–21 (2003).
19. Afuah, A., Tucci, C.L.: Internet business models and strategies: Text and cases. McGrawHill Int. Ed. Manag. Organ. Ser. 2, 358 (2000).
20. George, G., Bock, A.J.: The Business Model in Practice and its Implications for Entrepreneurship Research. Entrep. Theory Pract. 35, 83–111 (2011).
21. Zott, C., Amit, R.: Business model design: An activity system perspective. Long Range Plann. 43, 216–226 (2010).
22. Zott, C., Amit, R., Massa, L.: The business model: Recent developments and future research. J. Manage. 37, 1019–1042 (2011).
23. Gassmann, O., Frankenberger, K., Csik, M.: The St . Gallen Business Model Navigator. Int. J. Prod. Dev. 18, 249–273 (2013).
24. Amit, R., Zott, C.: Value creation in e-business. Strateg. Manag. J. 22, 493–520 (2001).
25. Pateli, A., Giaglis, G.: A Framework for Understanding and Analysig eBusiness Models. BLED 2003. (2003).
26. Solaimani, S.: a Qualitative Analysis on Business Model Implementation: a Design Case on a Dutch Bank. Euram 2013. (2013).
27. Osterwalder, A., Pigneur, Y.: Designing Business Models and Similar Strategic Objects : The Contribution of IS. J. Assoc. Inf. Syst. 14, 237–244 (2013).

28. Morris, M.H., Schindehutte, M., Richardson, J., Allen, J.: Is the business model a useful strategic concept? Conceptual, theoretical, and empirical insights. J. Small Bus. Strateg. 17, 27–50 (2006).
29. Demil, B., Lecocq, X.: Business model evolution: In search of dynamic consistency. Long Range Plann. 43, 227–246 (2010).
30. 30. Jung, H.: Allgemeine Betriebswirtschaftslehre. Oldenbourg Wissenschaftsverlag, (2006).
31. Amit, R., Zott, C.: Business Model Design: A Dynamic Capability Perspective. J. Manage. (2014).
32. Mintzberg, H.: The rise and fall of strategic planning. Pearson Education (2000).
33. Piccinini, E., Gregory, R.W., Hanelt, A., Kolbe, L.M.: Transforming Industrial Business : The Impact of Digital Transformation on Automotive Organizations. In: Thirty Sixth International Conference on Information Systems. pp. 1–20 (2015).
34. Johnson, M.W., Christensen, C.M., Kagermann, H.: Reinventing your business model, https://hbr.org/2008/12/reinventing-your-business-model. (Accessed 29.07.2016).
35. Pateli, A.G., Giaglis, G.M.: A Research Framework for Analyzing eBusiness Models. Eur. J. Inf. Syst. 13, 302–314 (2004).
36. Yin, R.K.: Case Study Research: Design and Methods. SAGE Publications (2003).
37. Meredith, J.: Building operations management theory through case and field research. J. Oper. Manag. 16, 441–454 (1998).
38. Urquhart, C.: Grounded theory for qualitative research: A practical guide. Sage (2012).
39. Eisert, U., Doll, J.: Business Model Based Management. 360° - Bus. Transform. J. November 2, 15–28 (2015).

Barriers to IoT Business Model Innovation

Dominik Bilgeri[1] and Felix Wortmann[2]

[1] ETH Zürich, D-MTEC, Zürich, Switzerland
dbilgeri@ethz.ch
[2] Universität St. Gallen, ITEM, St. Gallen, Switzerland
felix.wortmann@unisg.ch

Abstract. The Internet of Things (IoT), in which virtually all physical things become connected to the internet, promises enormous economic potential. The IoT might disrupt entire industries and it forces companies to rethink their current business activities. In light of these challenges, research on business model innovation (BMI) can offer promising insights. This research paper aims to contribute to the emerging BMI literature by identifying innovation barriers in an IoT context. 16 barriers are identified based on ten expert interviews that were conducted with employees from five multinational companies. The contributions of our study might lay a fruitful ground for future research.

Keywords: Business Model Innovation, Internet of Things, IoT BMI Barriers.

1 Introduction

General Electric (GE) has been well-known for selling industrial hardware and maintenance services. However, recently the company got under pressure in new fields of competition facing non-traditional competitors such as SAP, IBM as well as big data and analytics startups [7]. Instead of offering reliable industrial equipment, these new competitors shift the customer value proposition towards the Internet of Things (IoT) and "deriving new efficiencies and other benefits through advanced analytics and algorithms based on the data generated by that equipment" [7, p. 91]. To meet the new competition and to address the challenges arising from IoT solutions, GE is currently transforming its entire business model [7]. In this respect, GE faces severe challenges, similar to many traditional manufacturing companies across industries [3, 7].

Little is known about how the IoT will change business models and even less about what IoT specific barriers hamper business model innovation (BMI) [cf. 14]. In fact, first studies already investigate general barriers to BMI [3, 6]. In addition, there is a large literature stream on technical IoT challenges [e.g. 4, 12]. However, to the best of our knowledge, there are no empirical studies that investigate IoT-specific BMI barriers in detail. Previous studies on IoT BMI provide anecdotal evidence and do not base their findings on empirical data [cf. 13]. To shed more light on the depicted research gap, this study identifies barriers to IoT BMI and can be seen as a first explorative step towards a better understanding of IoT adoption. Thereby, the study focuses on large

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Bilgeri, D.; Wortmann, F. (2017): Barriers to IoT Business Model Innovation, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 987-990

multinational companies that conduct technology-driven BMI in an IoT context. More specifically, our research aims to address the following research question (RQ):

What are specific barriers to technology-driven IoT business model innovation?

2 Background

The *Internet of Things* describes a broader vision that all sorts of objects become smart, able to gather data and communicate both with each other and the internet [5]. Porter and Heppelmann argue that the key novelty of IoT solutions lies in "[the changed] nature of the 'things'" [13], including their connectivity and the digital services that they facilitate [5]. Such offerings, spanning the digital and the physical world, require a broad array of enabling technologies [12], which is reflected in various value creation layers inherent to IoT solutions [5].

Despite the lack of an agreed business model definition [15], a common high-level understanding emerged that business models explain a focal firms' value creation and value capture [17]. Taking a transformational perspective on business models is the essence of *business model innovation* [4]. In the light of a missing conceptualization, business model innovation is used as an umbrella term describing companies' efforts connected to "the search for new business logics of the firm and new ways to create and capture value for its stakeholders" [2, 9]. So far, little is known about BMI processes [14]. In fact, scholars agree that the related literature on new product development (NPD), with its much richer record, is most suitable as a starting point for BMI process research [3, 6, 14]. Latest research on NPD identifies four key innovation stages: idea generation and screening (ideas are gathered and the most promising opportunities selected); concept development and evaluation (business model components, such as revenue mechanics, are elaborated); technical implementation (technical realization of the offering and introduction of the business model throughout the organization) as well as commercialization (business model is scaled successfully in the market) [10].

The complexity of emerging IoT solutions brings along some severe *new challenges* [cf. 12]. A rich field of research exists, investigating IoT barriers from a very technical viewpoint. Among the most critical problems identified in this type of literature are a lack of protocol standardization, scalability limitations, energy supply and security issues [cf. 4, 12]. Besides this technical stream, various studies, including the seminal article of Chesbrough, investigate what might generally hinder organizations to innovate their business models [cf. 3]. Amit and Zott [1], for instance, identify the four characteristics novelty, lock-in, complementarities as well as efficiency and elaborate how these aspects might be contradicting to "traditional configurations of firm assets" [3, p. 358]. Several early studies also elaborate on more operational challenges managers face when attempting to innovate their companies' business models, such as overcoming internal resistance, financial hurdles, setup of value networks and the successful management of the applied implementation approach [6].

3 Methodology

Ten in-depth expert interviews were conducted with employees from five leading multinational corporations across the IoT ecosystem. The interviewee selection followed a heterogeneous purposive sample approach applying three pre-defined criteria [16]: (1) interviewees have either been actively involved in or closely guided IoT related BMI projects; (2) they possess more than two years of IoT BMI experience to ensure that they can sufficiently inform the research; (3) they are employees, partners or consultants to manufacturing companies across the IoT ecosystem. Interviews lasted between 30 and 60 minutes, followed the same case protocol, were audio-recorded and additional secondary data about the companies in general or mentioned IoT BMI projects in particular were collected for data triangulation purposes [16].

4 Findings and conclusion

The paper at hand identifies 16 barriers to IoT business model innovation, which are structured along four high-level innovation stages (cf. Table 1) [10].

Table 1. Identified IoT BMI barriers structured along stages adapted from Luchs et al. [10])

I. Idea generation and screening	*II. Concept development and evaluation*	*III. (Technical) implementation*	*IV. Commercialization*
1. Process ownership 2. Product-centric focus 3. Immediate profit thinking 4. Uncertainty	5. Innovation-driven R&D vs. reliability-focused IT 6. Collaborate with various partners 7. Data privacy 8. IoT value propositions	9. Clash of cultures 10. Data analysis 11. Legacy systems 12. Efficient operations	13. IoT revenue mechanics 14. After-sales commercialization 15. Customer analytics 16. Ambidexterity

Thus, this study can be seen as a first explorative step towards a better understanding of IoT adoption. The findings contribute to the ongoing debate on BMI by strengthening scholars' and practitioners' understanding of IoT specific hurdles to business model innovation. Some of the identified barriers build upon well-known BMI barriers, such as *legacy systems* or *ambidexterity* [8, 11]. But, when analyzed in an IoT context these barriers gain unexpected new facets and thus relevance. Others are IoT specific and have not yet been discussed in the BMI literature at all, including barriers such as *data analysis* or *customer analytics*. The results of this study should be assessed in the light of their limitations. One general limitation of qualitative research is generalization of results [16]. More specifically, our research is limited by the number and selection of interviews conducted. Further work in different empirical settings will be required to enhance the validity of the research. Taking these aspects into account, we hope that the contributions of our study lay a fruitful ground for future research. Promising research avenues might include a more in-depth analysis of the identified IoT BMI barriers with regard to their impact on BMI success.

The presented work is supported by the Bosch IoT Lab at the University of St. Gallen.

References

1. Amit, R., and Zott, C. 2001. "Value Creation in E-Business," Strategic Management Journal (22:6-7). pp. 493-520.
2. Casadesus-Masanell, R., and Ricart, J. E. 2010. "From strategy to business models and onto tactics," Long Range Planning (43:2), pp. 195-215.
3. Chesbrough, H. 2010. "Business model innovation: opportunities and barriers," Long Range Planning (43:2), pp. 354-363.
4. Cortimiglia, M., Frank, A., and Ghezzi, A. 2015. "Business model innovation and strategy making nexus: Evidence from a cross-industry mixed-methods study," R&D Management, doi: 10.1111/radm.12113.
5. Fleisch, E., Weinberger, M., and Wortmann, F. 2014. "Geschäftsmodelle Im Internet Der Dinge," HMD Praxis Der Wirtschaftsinformatik (51:6), pp. 812–826.
6. Frankenberger, K., Weiblen, T., Csik, M., and Gassmann, O. 2013. "The 4I-Framework of Business Model Innovation: An Analysis of the Process Phases and Challenges." International Journal of Product Development (18:3-4), pp. 249-273.
7. Iansiti, M., and Lakhani, K. R. 2014. "Digital Ubiquity: How Connections, Sensors, and Data Are Revolutionizing Business," Harvard Bus. Rev. (92:11), pp. 91-99.
8. Koh, S., Gunasekaran, A., and Goodman, T. 2011. "Drivers, barriers and critical success factors for ERPII implementation in supply chains: A critical analysis," Journal of Strategic Information Systems (20:4), pp. 385-402.
9. Landau, C., Karna, A., and Sailer, M. 2016. "Business model adaptation for emerging markets: a case study of a German automobile manufacturer in India," R&D Management, doi: 10.1111/radm.12201.
10. Luchs, M., Swan, K., and Creusen, M. 2015. "Perspective: A Review of Marketing Research on Product Design with Directions for Future Research," Journal of Product Innovation Management (33:3), pp. 320-341.
11. Markides, C. 2013. "Business model innovation: what can the ambidexterity literature teach us?," The Academy of Mgt Perspectives (27:4), pp. 313-323.
12. Mattern, F., and Floerkemeier, C. 2010. "From the Internet of Computers to the Internet of Things," in From active data management to event-based systems and more, K. Sachs, I. Petrov, and P. Guerrero (eds.), Berlin: Springer, pp. 242-259.
13. Porter, M. E., and Heppelmann, J. E. 2014. "How smart, connected products are transforming competition," Harvard Business Review (92:11), pp. 11-64.
14. Schneider, S., and Spieth, P. 2013. "Business model innovation: towards an integrated future research agenda," Intern. Journal of Innovation Mgt. (17:1), 1-34.
15. Teece, D. J. 2010. "Business Models, Business Strategy and Innovation," Long Range Planning (43:2-3), pp. 172–94.
16. Yin, R. 2013. Case Study Research: Design and Methods, Thousand Oaks: Sage.
17. Zott, C. and Amit, R. 2009. "Business model design: an activity system perspective," Long Range Planning (43:2), pp. 216-226.

Critical Success Factors of Digital Business Strategy

Friedrich Holotiuk[1] and Daniel Beimborn[1]

[1] Frankfurt School of Finance & Management, Frankfurt, Germany
{f.holotiuk,d.beimborn}@fs.de

Abstract. Digitalization does fundamentally impact firms' strategy development. With the fusion of IT and business strategy, Digital Business Strategy (DBS) creates the foundation for digital business models [1]. In this paper, we develop a DBS framework, based on a structured review of 21 industry reports. From this analysis, we yield 8 generic dimensions with a total of 40 critical success factors (CSFs) for DBS. The CSFs represent a rich set of actions specific to DBS and to the design of business models in the digital business environment. The discussion shows that academic research is lagging behind in contributing to DBS. Future research is suggested to further formalize the concept of DBS and to create a better understanding about how firms can successfully establish DBS.

Keywords: Digital Business Strategy, Business Model, IT Strategy, Industry Reports, Meta-Analysis

1 Introduction

Digital Business Strategy (DBS) is an emerging concept at the intersection of Information Systems (IS) and Strategic Management [2], which calls for contributions from academic research [3]. DBS describes the fusion of business and IT strategy [1] and the incorporation of digital technologies in business strategy [4]. With DBS, the separate and subordinate role of IT strategy to business strategy is given up for a joint approach to both, thereby leveraging internal, e.g. IT, instead of externally focused actions to create competitive advantage [2]. Recently, this trend has got into the focus of IS journals, like the MISQ special issue 2013 or the research agenda proposed in Business and Information Systems Engineering (BISE) in 2014 [5]. The combination of different scientific perspectives as intended by Veit et al. [5] is a cornerstone of DBS. However, guidelines for the development of DBS along with effective implications for the design of digital business models (BMs) are still scarce in the academic literature [6].

Contrary, DBS has found strong support in practice in form of numerous industry reports of research centers like the MIT Center for Digital Business [7], research firms like Gartner [8], technology advisory firms like Accenture [9] or Capgemini [10], and strategy consultancies such as McKinsey [11], BCG [12], and Bain [13]. In practice, the advancement of digital technologies and the rethinking of strategy are linked to "strategic principles" [14 pg. 5], a "digital strategy process" [9 pg. 10], or "strategic

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Holotiuk, F.; Beimborn, D. (2017): Critical Success Factors of Digital Business Strategy,
in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung
Wirtschaftsinformatik (WI 2017), St. Gallen, S. 991-1005

challenges" [15 pg. 9]. The use of digital technologies to "achieve strategic end" [16 pg. 6], leverage of data to create customer insight [8], and develop a digital mindset [17], are among the clearly identified actions of DBS in practice. Interlinks between practice and academia are observable with the new line of thinking on the role of IS executives, inspired by John F. Rockart, developing from "technically oriented" to "managerially oriented" [18 pg. 3]. IS executives are now developing more 'strategically oriented' skills along with greater digital literacy [19], as digital technologies become of strategic importance [20]. These technologies are summarized under digitalization, which stands for the emergence of "rapid Internet and technological advancements" [21 pg. 48] and "a sociotechnical process of applying digitizing techniques to a broader social and institutional contexts that render digital technologies infrastructural" [22 pg. 2]. As a result, strategic implications for the design of BMs from digitalization are observable, justifying the current research on DBS in academia and mainly practice. Nonetheless, good practices for implementing DBS or insights into strategic actions of DBS are few [6]. The development of the related academic research stream is still in its infancy. Hence, it is worthwhile to study what industry reports are suggesting firms to do in terms of Critical Success Factors (CSFs), i.e., in order to "ensure successful competitive performance" [23 pg. 12]. CSFs define a few actions and areas of a firm which are of critical importance to the firm's success [23]. They are not focused on achieving competitive advantage but lead to competitive failure if not met [24]. Hence, they can be used to answer currently open questions on effective strategy development of DBS and the successful translation of DBS into BMs. Moreover, CSFs help us to gain a better understanding of DBS in practice and provide helpful insights for academia. In the end, we deducted our research question directly from the outlined shortcomings of academic research as well as from the profound calls for new insights to DBS: *What are the CSFs of DBS in today's business environment?*

Consequently, our research is motivated by the quest for a new, holistic, approach to strategy development in the digital business environment and the identification of CSFs of DBS. Thus providing new academic input and the needed theoretical grounding. The goal is to derive a set of CSFs of DBS, which can be synthesized in a framework for DBS and the design of digital BMs. As academic research findings on DBS are scarce, this paper provides a review of, particularly, industry reports to include findings and advancements on DBS from practice. We consolidate the current state of industry research on the nature and success of DBS as well as contribute a framework for DBS, which is structured along 8 dimensions with 40 CSFs. The dimensions of the framework provide managerial implications for designing BMs in the digital business environment and outline future research avenues. Also, implications from strategy to BMs further to business processes can be derived [25].

The remainder is structured as follows: First, the background provides a literature summary of the origin of DBS and its implications for the design of BMs. Next, the research methodology explains the detailed procedures to extract CSFs of DBS via a structured analysis of industry reports. Afterward, findings are synthesized in the DBS framework and the CSFs are explained. Finally, the discussion puts the framework in perspective to the literature and highlights implications and limitations of our research.

2 Background and Related Research

IT has been on the management agenda for a long time, it was first focused on data processing and technical problems, developed to be more functional-oriented and focused on Management Information Systems, later focused on first strategic implications and the competitive use of IT, and last to the alignment of business and IT [26]. Consequently, IT and business strategy were characterized by a distinction, such as how IT could support to achieve business goals. This distinction has been coined 'strategic alignment', describing the parallel development of IT and business strategy as well as the support function of IT to business [27]. As a result, an "aligned but essentially subordinate" [1 pg. 472] role of IT to business strategy dominates strategy development. This differentiation is inadequate for a time where digital technologies are disrupting products, industries, and markets. Where firms have to entrench digital capabilities to stay competitive. Hence, with digitalization a gap between what is required and the traditional thinking of strategic alignment developed.

DBS aims at addressing this gap by taking a broader and more inclusive view on IT strategy and recognizing changes due to the digitalization. Statements, like "digitalization transforms our society like only the industrial revolution did before in recent history" [28 pg. 10] and society is undergoing "remarkable change because of digital technology" [29 pg. 734] support the ongoing transformation. The fusion of the digital and physical worlds fundamentally impacts firms' operations and industry boundaries. In this transformation, it is not a matter of whether or not firms will be digital, but a matter of how and when they become digital [9]. Strategy plays a superior role in this transformation; Catlin et al. [30] show that strategy-related factors cause the biggest variance between 'digital leaders' and average performers when comparing competitiveness. However, the transformation of strategy is lagging quite behind the transformation of the business environment.

As a result of digitalization, IT strategy is seen today "as essential to the framing of overall business strategy, that is, a fusion of IT and business strategy" [2 pg. 513]. The implied constant interlocking of business and IT strategy with DBS can deliver competitive advantage [31,32] and underlines the contrast to the traditional view of IT strategy. However, the current state of research on DBS unveils blank spots in IS as well as Strategic Management with only a few academic articles having been published. While IS research has analyzed how digital technologies are influencing a firm's strategy, structures, and processes [32,33] and how business value is shifting [34], the fusion of business and IT strategy has not yet been a clear focus – the question of "how to build a competitive advantage in turbulence with digital IT systems" [35 pg. 443] remains.

However, it should not be neglected that, academia needs naturally more time than practice to establish a new research stream. Consequently, the overall research on DBS has remained largely atheoretical and dominated by industry reports [16] and practitioner guidelines [36]. Research in practice shows that "conventional business strategy is not the best fit for meeting the demands of digital growth" [9 pg. 8] and affirms that "companies [are] competing in a fundamentally new way" [14 pg. 2], due to digitalization. Practice research recognizes that DBS "it is not about changing the way we do technology, but changing the way we do business" [7 pg. 22] by rethinking strategy.

Although implications on strategy are discussed, BM-related considerations are few, but should be considered, as these two concepts, strategy and BM, although they are distinct, are connected [5]. Strategy is about “deliberately choosing a different set of activities to deliver a unique mix of value” [37 pg. 64] whereas the BM is “an abstract representation of an organization […], of all core interrelated architectural, co-operational, and financial arrangements […], as well as core products and/or services […] needed to achieve strategic goals” [25 pg. 8]. BMs are one way to translate strategic actions from DBS into value architectures for the business [38]. Hence, they have undergone tremendous change as value creation has shifted, due to redefined product and service offerings [39]. New BMs are needed as traditional products and industries are disrupted by digitalization. “The digitalization of the book is fundamentally reshaping the structure that has underpinned the book publishing for 200 years […], changing the very idea of books” [29 pg. 724]; illustrates how value shifts away from products, where it is unclear if industries or markets stay the same or have to adapt.

So far, there is little research about how firms design DBS to develop new BMs as a way to create opportunities and stand out from the industry norm [2]. Nevertheless, the expectations of the research community are high, such as the hope for “DBS to explain why some [strategic implications of digitalization on the business environment] […] are observed instead of others” [40 pg. 553]. Our research connects the consequences of digitalization with implications on strategy as we see “digital technologies are fundamentally reshaping traditional business processes” [1 pg. 472], provides insights into DBS in practice from industry reports, and synthesizes findings into a DBS framework.

3 Research Methodology

Insights from industry reports with a focus on strategic aspects due to digitalization promise to help to develop a better understanding of DBS in academia. However, so far the reports fail to put their research within a theoretical framework and to provide significant insights for academic research, which we aim to overcome. The following review and analysis of industry reports will derive CSFs of DBS, which deliver the input from practice for the development of the DBS framework. Our approach extends current academic research with a profound understanding of the nature of the concept in practice, which is specifically suggested for DBS research [1] and also applied for similar emerging IS concepts [41]. For the review of industry reports, we followed the guidelines for scientific literature reviews as specified by Webster and Watson [42]:

Data collection started with the identification of relevant industry reports. Reports from, among others, analyst and consulting firms were identified by an online search via public internet search engines for the term ‘digital business strategy’ and slight variations of it (e.g. digitalization strategy). The search was restricted to reports published between January 2011 and July 2015 for the following reasons: First, insights and themes in industry reports are extremely fast-moving and the publishing rate is high, with around five new reports per month. Second, the strategy focus of digitalization is new and it was not until recently that reports focused on strategic implications and DBS [43]. Third, our research should build on a fairly mature stock of insights, which is rather true for more recent reports, as they are building on the knowledge gained from previous ones. Next, from the online search results, the first 200 hits were reviewed to find out whether they included an industry report relevant to our research goal. Reports had to specifically address strategic aspects in response to digitalization. After deletion of duplicates, the search delivered a total of 64 industry reports.

These reports were then scanned for their strategic implications (e.g. the design of new organizational structures, the long-term orientation of the firm), relevance for strategy in general (the development of new strategies, involvement in the strategy making), as well as usefulness to develop a framework for DBS (e.g. first indications for possible dimensions). Industry reports fulfilling these requirements hold the most potential to deliver input for the framework and were therefore included. We included, for example, a report focusing on strategy over technology [16] or a study on growth strategies [9].

Industry reports not fulfilling these specifications were excluded when they met at least one of the following criteria: (1) focusing on only one singular organizational function like sales or digital marketing; (2) focusing purely on new leadership roles (e.g. only discussing responsibilities of employees or managers); (3) associated with DBS but not providing any strategic implications like an analysis of German ‘Industry 4.0’ initiatives in the production; (4) too small deltas to other reports (due to subsequent reports from the same authors/institutions) (5) specific industry focus and no implications for strategy (e.g. focused on technological vision). Finally, 21 industry reports were selected as input for the subsequent analysis [7–17,20,30,36,43–49].

For analyzing these 21 reports, we applied the ‘critical success factor method’ [23]. It aims to identify and determine information and actions (CSFs) which are most needed to reach a defined outcome or goal [23]. Although the original version of the CSF method builds on interviews with top-level executives, it could be modified for our research by replacing the interview data source with the identified industry reports. We used a simplified coding method to identify the specific CSFs in each report [50]. The original aim to make CSFs explicit and encapsulate them to management priorities, allowing management to act on a more knowledgeable and [23], remained true.

Subsequently, the coding of the 21 industry reports assisted our research by “organize[ing] and make[ing] sense of the qualitative data” [50 pg. 152]. To obtain our framework we followed a staged process: First, we coded CSFs which explicitly focus on strategic actions undertaken by firms or identify potential for firms to improve strategy development to address digitalization. Identified CSFs were recorded in an Excel spreadsheet. The coding delivered an initial, cumulative, and duplicate-free set of 89 CSFs, which occurred in the sample. The second step was a repeated analysis of the 21 reports for the mentioning of the previously determined set of 89 CSFs, which created a result matrix with information on which CSF was mentioned in which report. As a result, the analysis delivered a total of 500 entries with an average of 23.8 different CSFs per report and 5.6 reports per CSF. Counts for each CSF ranged from 1, the lowest (CSF found in only one report), to 21, the highest (CSF found in all reports). This comprehensive set of weighted CSFs provided rich insights but was still distorted by CSFs that were raised to be ‘critical’ by only very few reports. Therefore, we followed Sproull [51] and reduced the set of CSFs to only those that were identified in more than 15% of the reports (i.e. 3 reports) to gain a more robust set of CSFs. The eliminated 30% of

CSFs explained less than 11%; contrary, the final set of 62 CSFs explain 90% of all entries. Ultimately, we consolidated CSFs with overlapping content as well as grouped similar CSFs together. In the end, 40 CSFs provided the final input for the framework.

Finally, we screened the CSFs to identify dominant themes. Applying an iterative procedure starting with rather broad themes and trying to further condense them to most relevant, led to the creation of 8 dimensions forming our DBS framework. These dominant themes are comparable to themes in academic research, as used by Kohli and Grover in their article, describing themes as “not mutually exclusive, nor are they exhaustive; rather they are meant to initiate a discussion of how the IT community must rethink […] and expand the agenda for research.” [34 pg. 28]. Analogously, the themes were used to structure the CSFs in practice and the dimensions of our DBS framework.

4 Findings

The layout of the DBS framework (Figure 1) is inspired by similar frameworks in adjacent research, like a framework synthesizing the literature on agile manufacturing [52], and is oriented towards the ‘business model canvas’, which structures relevant areas for new BMs [53], to provide the same clarity and guidance for DBS and to support the translation of strategy into BM by a corresponding structuration. Our framework comprises the 8 dimensions, identified as dominant themes, and the 40 CSFs.

Figure 1. The DBS framework

The first two dimensions, <u>*Sales and Customer Experience*</u> and <u>*Organization*</u>, are the two largest subsets of CSFs; hence and due to their fundamental role, they are pictured as the two ‘pillars’ of the framework. All other dimensions are placed in a descending order with regard to their strength (i.e. count of entries) between the pillars. Table 1 shows all dimensions, the assigned CSFs, and the count of entries. The CSFs provide actions for DBS development and the design of new BMs. They are introduced in detail in the following, while only explaining the strongest CSFs for each dimension:

<u>*Sales and Customer Experience*</u> focuses on *seamless integrated offline (physical) and online (digital) channels* [8,16]. The seamless customer experience across all channels with an integration of channels is a decisive action [13,49]. Mobile, as the most significant sales channels in future, requires a coherent presence across all channels. DBS builds on blending physical and digital worlds by a greater integration of online and offline experiences. Experiences around the product, mostly digital and intangible, are superior to the physical utilization of the product [13,36].

Table 1. The dimensions and CSFs as identified by the analysis of industry reports

#	*Count*	*Dimensions / CSFs*	*#*	*Count*	*Dimensions / CSFs*
	84	*Sales and Customer Experience*		*76*	*Organization*
1	20	Seamlessly integrated offline(physical) and online (digital) channels	7	15	Agility to reallocate resources and reorganize rapidly
2	15	Digitalization of customer interaction and products&services	8	13	Change management for radical and rapid change
3	14	Analytics to customize and create products&services	9	12	Multi-level and multi-speed organization for faster reaction
4	13	Direct contact for customer centricity			
5	12	Customer integration with open innovation	10	11	Organizational alignment towards digital
6	10	Outstanding customer experience and satisfaction	11	10	Long-term orientation but short, intense sprints to change
			12	8	Organizational separation → Spin-off
			13	7	Lean decision-making
	66	*Culture and Leadership*		*57*	*Capabilities and HR Competencies*
14	15	Create and foster digital mindset with a digital agenda	20	13	Capability to reinvent value chain and to challenge status quo
15	15	Common set of values with digital as value creation			
16	12	Accept failure and encourage new to grow success	21	11	Digital skills, know-how, and talent
17	10	Innovation and adaptive culture with evolvable goals	22	10	Capability to design new business models
18	8	Commitment to transformation in strategy and culture	23	9	New assets and capabilities
19	6	Rethinking of C-level roles (CDO, CIO)	24	7	Leaders have to identify new HR potentials
			25	7	Acquire, retain, and attract new talents
	56	*Foresight and Vision*		*48*	*Data and IT*
26	13	Establish a clear vision with future positioning	31	19	Use data and information from central source
27	13	Tight feedback loops and aspiration to improvements			
28	12	Foster faster innovation / rapid prototyping	32	11	Fundamentally different role of IT with two-speed IT
29	10	Look what is laying left and right			
30	8	Bold experimentation	33	10	Real-time and large-scale data processing
			34	8	Modular IT platform
	31	*Operations*		*25*	*Partners*
35	17	Data-driven and digitally automated process	39	14	Network effects with open systems and partner integration
36	5	Not just business but operating models change			
37	5	Blending human and digital resources	40	11	External partners
38	4	Provide financial resources			

Augmented reality, customer-focused technology, and digital customer decision journeys lead to a *digitalization of customer interaction and products&services*, allowing firms to develop data-rich insights about customers [16]. As products and services are being digitalized firms are required to enhance them with digital extensions like a permanent access to them via a digital channel. As a result, BMs have to allow for every (physical) product being augmented with a digital service [8]. The increased availability of customer makes *analytics to customize and create products&services* an absolute necessity with DBS where customer messages are individualized to preferences and enriched with contextual data. Products are constantly adjusted to better reflect the customer needs. With *direct contact for customer centricity*, firms build deeper connections between brands and customers as well as execute customer care from all business units via digital channels [36]. Firms have to avoid third parties and disintermediate customer relationships to establish a long-lasting tangible and emotionally affected brand connection [15]. Firms take on an 'outside-in' perspective, internalize the customer viewpoint, and customers become the central hub of digital service delivery [49].

Organization focuses on *agility to reallocate resources and reorganize rapidly*. Agility is an integral part, or, 'in the DNA' of DBS, requiring firms to self-tune the organization to changing circumstances. DBS builds on organizational agility to allow for fast adaptation [16] by collaboration across foregone organizational boundaries. Organizational liquidity allows firms to shift businesses as customer needs shift, requiring new BMs, agile operating models, and the ability to scale fast and learn quickly [43].

Proper *change management for radical and rapid change* prepares employees for imminent changes and sets the organization on a new course. Story-telling can be used to convey the ‘digital story’ and to gain employee buy-in, stimulate pride, and entrench the digital transformation [16]. A *multi-level and multi-speed organization for faster reaction* is needed to enable ambidextrous organizations which renovate the core and innovating the outside. As a result, firms are able to quickly respond to customers’ demand while balancing internal constraints with needed speed for faster reaction. *Organizational alignment towards digital* is used to establish a dedicated team to support digital business opportunities and build up digital governance to align fragmented digital activities [30]. Digital transformations by individual units need to be aligned to the organization, culture, and technology of the firm [47]. *Organizational separation, such as spin-offs,* enable digitally centered sub-businesses with separate competitive advantage. Corporate venturing is a catalyst for digital platforms and growth of divisions without constraints by utilizing variable business architectures [12,17].

Next, *Culture and Leadership* aims to *create and foster a digital mindset with a digital agenda* where the culture takes on an exploratory and adaptive character [14], is open for change, and “conducive to the digital transformation” [16 pg. 9]. Collaboration and cross-functional work are encouraged to generate new ideas and drive innovation. Firms are breaking free of silo-thinking [30]. With it, firms have to establish *a common set of values with digital as value creation* and integrate digital technologies in the transformation as well as the way people work. Leaders have to entrench digital values for the culture, such as forward-thinking, openness, technology acceptance, entrepreneurial spirit, and a startup way of working [43]. The value of digital innovation has to be understood, recognized, and cherished to be successful [16]. This goes along with a new culture which supports to *accept failure and encourage new to grow success.* It is necessary to establish a common appreciation that risk taking involves failure and failure is embraced “as a prerequisite for success” [16 pg. 4]. The culture must encourage risk-taking and tolerate failures to succeed [14]. A *commitment to transformation in strategy and culture* is needed by the leadership team to set DBS on the right course and lead the digital transformation by example from the top of the organization [7,16]. It is important to understand that leaders drive the transformation, address technologies that bring change, and trigger the connected change of culture.

The *Capabilities and HR Competencies* dimension strongly builds on a *capability to reinvent value chain and challenge status quo* where employees identify where value is now and in the future as well as lift value to the next level by moving it from the traditional world of value chains to the world of platforms, ecosystems, and stacks [12]. Disaggregated value chains are caused by a reduction of transaction costs and firms need to analyze the value chains to detect points which are best for possible digitalization [13]. The need for *digital skills, know-how, and talent* stands on top of the list for almost all firms [47]. In order to design and execute DBS, firms have to train employees for the needed digital skills, align incentive systems, and provide financial resources for human capital development [7,16]. Beyond the clear technical IT skills, also non-IT skills, such as visioning, collaboration, and organizational change management, are required with DBS. New talents are attracted by a commitment to the digital transformation, leaders' digital literacy, and a firm-wide understanding of the power of digital technologies [16,43]. In particular, the *capability to design new business models* becomes a strategic capability with DBS. As BMs change from 'inside-out' to 'outside-in', they become more customer focused. "This shift is the essence of adopting a digital BM" [9 pg. 6]. Firms innovate their BMs more often and incorporate small-scale as well as fundamental disruptions of traditional BMs [16]. Hence, BMs are shorter lasting and adapted more frequently when aligned to DBS [8]. The knowledge-driven digital business environment requires *new assets and capabilities*, such as, in the form of state-of-the-art infrastructure to enable the digitalization of products, as well as digital capabilities to design digital product extensions [49]. Firms digitally transform their core capabilities by building up complex and cross-functional combinations of assets and capabilities in their BM, e.g. people, processes, and expertise, to withstand competition [17]. However, not everybody is able to learn required digital skills, so *leaders have to identify new HR potentials*. Hence, the assessment of required skills and the identification of people, which can be trained to support the transformation, are needed [8,30].

<u>*Foresight and Vision*</u> is about *establishing a clear vision with future positioning* for the digital transformation [20]. With DBS, firms are required to have a transformative vision of the future and still provide the needed clarity and to achieve it [7]. The vision is characterized by *tight feedback loops and aspiration to improvements*, where firms learn from customer, employee, and partner interactions to further develop the vision and, hence, update services and products. Frequent feedback loops and the reaction to feedback are an iterative process which determines the success with DBS [47]. *Fostering faster innovation/ rapid prototyping* enables 'learn-track-react' behavior and 'test-and-learn' approaches, where firms model new products quickly, put them into the market promptly, and test them constantly [49]. For the continuous delivery of new products, minimum viable products are sufficient which increase the speed of product development [49] to allow firms to "fail, fast, and inexpensive" [8 pg. 24]. As the environment is changing quickly, firms need to *look what is laying left and right*. Scanning the environment helps to identify digital opportunities, disruptions, and potential threats [30]. Sensing and anticipating of technology-driven transformations are needed [12] as traditional industry barriers disappear and allow asymmetrical rivals and unlike allies.

The <u>*Data and IT*</u> dimension puts the CSF to *use data and information from a central source* in focus. It is one of the most often occurring CSFs in our findings, which underlines that (big) data analytics and sense-making of structured and unstructured, as well as inside and outside data from different sources channeled in a central data source, are vital [30]. DBS leverages data and analytical methods to use data-driven decision making, make data-assisted economic decisions, learn about the customer, and turn data into insights. Data is a competitive advantage as information is at the core of BMs in the digital business environment. As a result, the *fundamentally different role of IT with two-speed IT* is no longer about enabling, but creating the business. IT takes on a new thinking by harnessing digital technologies to create business value [20]. *Two-speed IT* lets firms operate IT at two different speeds with rapid results, high reaction times, and extreme flexibility towards the customer facing side, and a strong internal backbone on the other [30,47]. This bimodal approach supports digitalization with rapid front-end changes while fulfilling backend requirements. Moreover, *real-time and large-scale data processing* is a key action allowing firms to "track and communicate digital key performance indicators frequently” [30 pg. 11]. Additionally, real-time contextual information and analytics are combined to rapidly develop actionable insights from data for DBS [49]. A *modular IT platform* with “agile technology delivery skills” [30 pg. 5] builds on speedy but flexible services, and integrated functions with shared solutions. Firms utilize an orchestration of services along with a ‘continuous delivery model’, which allows them to release and iterate quickly [48].

Operations is infused with data to create *data-driven and digitally automated process* for higher automation. This allows supply chains to react quickly and anticipate customer demand. The further automated handling of services and completely automated customer interactions in BMs, such as Zipcar or car2go, increase speed and efficiency [11]. *Blending human and digital resources* creates human-centric designs for businesses with individual solutions by interlocked human and digital channels [49]. DBS focuses on efficiency in the interaction between people and technology, with digital technologies adding velocity to processes and services.

The main goal of the *Partners* dimension is to utilize *network effects with open systems and partner integration*. There is an increase in value with each new customer added, hence digitally-enabled firms tend to form natural monopolies and create “winner-takes-all dynamics”, e.g. Google and Facebook [15]. Firms have to open their services and products to a community as well as allow for an easy integration of new connected devices, objects, and people via open standards, allowing products in a network to be more powerful. With *external partners,* firms form strong and collaboratively partnerships. Additionally, extensive external orientation supports learning and innovation [8]. The collaboration goes beyond boundaries of the firm and extends to customers, technology providers, and suppliers. Firms allow partners to collaborate for specialized expertise [49] and utilize partnerships for specific innovations.

In summary, all 40 CSFs allow better strategy development and help to advance the understanding of DBS, but only some have received academic mentioning.

5 Discussion

We advance the literature on business and IT strategy by identifying CSFs of the content and development of Digital Business Strategy. Specifically, we presented a comprehensive overview of CSFs, as raised by industry reports, consolidated in our DBS framework. The framework provides actions for the development of DBS with 40 CSFs, divided into 8 dimensions. As the framework is derived from industry research, its CSFs should be discussed by reflecting it with scientific work, where available.

First, the academic literature exhibits a strong focus on the customer side of DBS and on building customer-centric organizations [4], which supports the strong customer orientation in our framework, as represented by the pillar ‘sales and customer experience’. We observe a growing number of digital BMs, redefining customer interfaces with digital technologies and putting customer service capabilities to the next level. As a consequence, BMs have to ‘integrate physical and digital worlds’ by blurring boundaries between online shops and physical stores (e.g., order online and pick up offline).

Furthermore, as IT systems become core elements of firms, there is evidence of firms 'digitalizing products&services' for competitive advantage [29,54]. With it, the 'digitalization of customer interaction' leverages customer data to improve in-store customer satisfaction by working with value-adding service from online, like access to customer profiles to provide individualized services. Recent examples are the BMs in the sports gear industry, which focus on integrated fitness trackers for all apparels.

Next, there is a growing body of literature on new leadership roles due to the digitalization. For instance, Bennis, in line with the 'culture and leadership' dimension of our framework, focuses on the change of leadership roles with DBS as he discusses how leaders must "understand the power" of digitalization [19 pg. 635]. In practice, as in our framework, the role of the Chief Digital Officer (CDO) pays respect to this development and ensures the needed attention. As a consequence of DBS, we see the appointment of CDOs to head new BMs and addressing digital services. Contrary, academic research remains at stressing skills like transparency and adaptive capacity [19].

Nonetheless, changes due to DBS should not stop at the top level, but have to be translated further into the entire organization. Hence, firms have to establish 'organizational alignment towards digital', which has not been mentioned in the DBS literature, yet. The scarce literature on DBS stresses to increase the firm's strategic capability to adapt to dynamic changes in the business environment [1]. Indeed, 'agility' becomes a strategic action and goal, as we were able to identify it as a CSFs of our framework.

Further, developing two organizational speeds is a way for creating organizational agility; it is equally discussed in practice as well as academia. The so-called 'multi-speed organization' is a CSF and fundamental tension of DBS. It is supported by Woodard et al. [40], who stress dynamic adaption to changing market conditions while providing a stable environment for value generation, and the recent bimodal IT concept, describing "different architectures, processes, and organization parts" [41 pg. 1418].

Also, other CSFs, outlined in our framework, yield opportunities for digital BMs like to 'use of data and information' (e.g. 'real-time and large-scale data processing'), or integrating niche players to access highly-specialized services in value networks.

In conclusion, academic articles on DBS fall short with regard to some aspects of the DBS framework but show first insights into DBS from academia. The discussed aspects of the DBS literature present only a subset of the CSFs identified in our framework. The DBS framework with its 8 dimensions and 40 CSFs, presenting the nature of DBS in practice research, goes beyond currently scientifically addressed topics, offers a broad perspective on the strategic implications due to digitalization, and provides guidelines for strategy development and BM design. Therefore, DBS focuses on the utilization of digital technologies and how they can be applied to create business value, hence, changes the way of business and revenues, ultimately leading to digital BMs [5].

Limitations and future research: Our research is limited by a focus on DBS relevant aspects, but the high number of CSFs shows the different facets of the concept and the diverse research areas. Hence, a deeper analysis of singular CSFs will be promising. Methodologically, the combination of a literature review approach and a modified critical success factor method to analyze industry reports is quite unique to IS research, so far; accordingly, the value and plausibility of such an approach might be less clear as for well-established research methods. Furthermore, the selection of industry reports is intended to provide a clear strategic focus for the DBS framework. Still, selection criteria can vary and are not free from subjective influences. The same applies to the subsequent elimination of minor CSFs, which was chosen according to the research set-up [51]. The CSFs were grouped into 8 different dimensions, using the dominant themes identified by the researchers. The themes are not free of overlap nor indisputable, a deeper analysis should look at relationships among them (e.g., interaction effects or causalities, preconditions etc.), since their impact on long-term firm success is of course highly inter-related (e.g., the tight interconnection between the dimensions of organization and culture). Further, the dimensions could be validated by a Delphi study, which could also refine a ranking of the most important CSFs, not provided by our framework so far. Hence, future research could add to our research with more quantitative measurement scales and statistical metrics for the CSFs. Finally, two limitations were identified by interviewees[1]: firstly, the framework lacks a finance dimension, which would pay respect to the huge investments firms make to advance the digital transformation; secondly, a dimension considering the influence of regulation on emerging topics and developments in the digital business environment is missing.

In summary, the understanding of adequate strategy development is essential to design BMs [25]. As we know that BMs are an extension of DBS, we discuss how BMs can add causal relations between the components of DBS to generate value.

6 Conclusion

DBS, an emerging concept for strategy development, has caught high attention by practitioners and scholars. It presents a further advancement of IT strategy and the traditional alignment view. Our research explores the current state of industry research on DBS, identifies CSFs of DBS, and synthesizes them into a framework. Our framework addresses outlined challenges of traditional strategy development, where IT strategy is "aligned but essentially subordinate to business strategy" [1 pg. 472], by presenting 40 CSFs. Led by these CSFs, our framework allows better strategy development and a more integrated approach to the challenges raised by digitalization. The dimensions of the framework are stable as they are conceptually rooted as well as remain generic to allow for changes of single CSFs over time.

DBS puts firms in a superior position, when it comes to the digital transformation and evolving to digital maturity. Evidence for the success of DBS is found since the majority (~ 80%) of digitally mature firms have a DBS; by contrast, among digital immature firms only 15% do [16]. By translating DBS into digital BMs above-average results regarding revenue [9], cost efficiency, and capital productivity can be expected. The CSFs of our framework will help practitioners advance DBS and present starting points for further research, as outlined above. In the end, our framework enriches the body of knowledge on DBS, but shows also that more insights are required.

[1] We started discussing the findings of our analysis with experts from the industry who brought up these thoughts. We thank for their intellectual support.

References

1. Bharadwaj, A., El Sawy, O.A., Pavlou, P.A. & Venkatraman, N., 2013. Digital Business Strategy: Toward a Next Generation of Insights. *MIS Quarterly*, 37(2), pp.471–482.
2. Mithas, S., Tafti, A. & Mitchell, W., 2013. How a Firm's Competitive Environment and Digital Strategic Posture Influence Digital Business Strategy. *MIS Quarterly*, 37(2), pp.511–536.
3. Bharadwaj, A., El Sawy, O.A., Pavlou, P.A. & Venkatraman, N., 2013. Visions and Voices on Emerging Challenges in Digital Business Strategy. *MIS Quarterly*, 37(2), pp.633–635.
4. Setia, P., Venkatesh, V. & Joglekar, S., 2013. Leveraging Digital Technologies: How Information Quality Leads to Localized Capabilities and Customer Service Performance. *MIS Quarterly*, 37(2), pp.565–590.
5. Veit, D., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Kundisch, D., Leimeister, J., Loos, P. & Spann, M., 2014. Business Models. *Business & Information Systems Engineering*, 6(1), pp.45–53.
6. Sia, K.S., Soh, C. & Weill, P., 2016. How DBS Bank Pursued a Digital Business Strategy. *MIS Quarterly Executive*, 15(2), pp.105–121.
7. Westerman, G., Calméjane, C., Bonnet, D., Ferraris, P. & McAfee, A., 2011. Digital Transformation: A Roadmap for Billion-Dollar Organizations. *Accenture*. Available at: https://www.capgemini.com/resources/digital-transformation-a-roadmap-for-billiondollar-organizations [Accessed March 22, 2016].
8. Plummer, D.C., Fiering, L., Dulaney, K., McGuire, M., Da Rold, C., Sarner, A., Maurer, W., Karamouzis, F., Lopez, J., Handler, R.A., Frank, A., Olding, E., McIntyre, A., Short, J., Shanler, M., Jivan, R., Taylor, B., Polk, J., Sorofman, J., Drobik, A., Perkins, E. & Welch, K., 2014. Top 10 Strategic Predictions for 2015 and Beyond: Digital Business Is Driving "Big Change." *Gartner*. Available at: http://www.gartnerinfo.com/exp /top_10_strategic_predictions_269904.pdf [Accessed April 10, 2016].
9. McManus, R. & McDonald, M., 2015. Growth Strategies for a Digital World. *Accenture*. Available at: https://www.accenture.com/ca-en/insight-growth-strategies-digital-world.aspx [Accessed April 10, 2016].
10. McGrath, R.G., Lemoine, P., Botsman, R., Solis, B., Cohen, D., Klein, S., de Bono, C. & Netessine, S., 2015. Digital Transformation Review. *Capgemini Consulting*. Available at: https://www.capgemini-consulting.com/sites/default/files/annual-report/983096/pdf/Digital_Transformation_Review_7.pdf [Accessed April 13, 2016].
11. Edelman, D. & Dörner, K., 2015. What "Digital" Really Means. *McKinsey*. Available at: http://www.mckinsey.com/insights/high_tech_telecoms_internet/what_digital_really_means?cid=digital-eml-alt-mip-mck-oth-1507 [Accessed March 21, 2016].
12. Evan, P. & Forth, P., 2015. Borges' Map: Navigating a World of Digital Disruption. *BCG*. Available at: http://digitaldisrupt.bcgperspectives.com/ [Accessed March 19, 2016].
13. Rigby, D.K. & Tager, S., 2014. Leading a Digical Transformation. *Bain*. Available at: http://www.bain.com/publications/articles/leading-a-digical-transformation.aspx [Accessed April 10, 2016].
14. Reeves, M., Nicol, R., Venema, T. & Wittenburg, G., 2014. The Evolvable Enterprise: Lessons from the New Technology Giants. *Bosten Consulting Group*. Available at: https://www.bcgperspectives.com/content/articles/future_strategy_business_unit_strategy_evolvable_enterprise_lessons_new_technology_giants/ [Accessed March 22, 2016].
15. Hirt, M. & Willmott, P., 2014. Strategic Principles for Competing in the Digital Age. *McKinsey*. Available at: http://www.mckinsey.com/insights/strategy/strategic_principles_for_competing_in_the_digital_age [Accessed March 20, 2016].

16. Kane, G.C., Palmer, D., Philips Nguyen, A., Kiron, D. & Buckley, N., 2015. Strategy, Not Technology, Drives Digital Transformation. *MIT and Deloitte*. Available at: http://sloanreview.mit.edu/projects/strategy-drives-digital-transformation/ [Accessed March 19, 2016].
17. Mainardi, C. & Vollmer, C.A.H., 2015. How Digital Leaders Outperform Their Peers. *Strategy&*. Available at: http://www.strategy-business.com/blog/How-Digital-Leaders-Outperform-Their-Peers?gko=324ad [Accessed April 6, 2016].
18. Rockart, J.F., 1982. The Changing Role of the Information Systems Executive: A Critical Success Factors Prespective. *Sloan Management Review*, 24(1), pp.3–13.
19. Bennis, W., 2013. Leadership in a Digital World: Embracing Transparency and Adaptive Capacity. *MIS Quarterly*, 37(2), pp.635–636.
20. Sahni, S. & Cobain, C., 2015. The New IT Horizon. *OliverWyman*. Available at: http://www.oliverwyman.com/insights/publications/2015/mar/the-new-it-horizon.html#.VaUeoPn5-5I [Accessed April 5, 2016].
21. Lerner, S., 2015. Digital Business Strategy. *Touro Accounting & Business Journal*, 2015(Spring), pp.49–52.
22. Tilson, D., Lyytinen, K. & Sørensen, C., 2010. Digital Infrastructures: The Missing IS Research Agenda. *Information Systems Research*, 21(4), pp.748–759.
23. Rockart, J.F., 1981. A Primer on Critical Success Factors. *Sloan Management Review*, 12(3), pp.7–15.
24. Freund, Y.P., 2002. Critical Success Factors. *Planning Review*, 16(4), pp.20–23.
25. Al-Debei, M.M., El-Haddadeh, R. & Avison, D., 2008. Defining the Business Model in the New World of Digital Business. In *AMCIS 2008 Proceedings*. pp. 1–11.
26. Teubner, R.A., 2013. Information Systems Strategy. *Business & Information Systems Engineering*, 5(4), pp.243–257.
27. Henderson, J.C. & Venkatraman, N., 1993. Strategic Alignment: Leveraging Information Technology for Transforming Organizations. *IBM Systems Journal*, 38(1), pp.472–484.
28. Vor dem Esche, J. & Hennig-Thurau, T., 2014. German Digitalization Consumer Report 2014. *Roland Berger*. Available at: http://www.rolandberger.com/media/pdf/Roland_Berger_German_Digitalization_Consumer_Report_20140718.pdf [Accessed May 10, 2016].
29. Yoo, Y., Henfridsson, O. & Lyytinen, K., 2010. The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research. *Information Systems Research*, 21(4), pp.724–735.
30. Catlin, T., Scanlan, J. & Willmott, P., 2015. Raising Your Digital Quotient. *McKinsey*. Available at: http://www.mckinsey.com/insights/strategy/raising_your_digital_quotient [Accessed April 10, 2016].
31. Mithas, S. & Lucas Jr, H.C., 2010. What Is Your Digital Business Strategy? *IT Professional*, 12(6), pp.4–6.
32. Prahalad, C.K. & Krishnan, M.S., 2002. The Dynamic Synchronization of Strategy and Information Technology. *MIT Sloan Management Review*, 43(4), pp.24–33.
33. Sambamurthy, V., Bharadwaj, A. & Grover, V., 2003. Shaping Agility through Digital Options: Reconceptualizing the Role of Information Technology in Contemporary Firms. *MIS Quarterly*, 27(2), pp.237–263.
34. Kohli, R. & Grover, V., 2008. Business Value of IT: An Essay on Expanding Research Directions to Keep up with the Times. *Journal of the Association of Information Systems*, 9(1), pp.23–39.
35. Pavlou, P.A. & El Sawy, O.A., 2010. The "Third Hand": IT-Enabled Competitive Advantage in Turbulence Through Improvisational Capabilities. *Information Systems Research*, 21(3), pp.443–471.

36. Penkert, A., 2015. He Who Serves, Wins. *Detecon*. Available at: https://www.detecon.com/us/en/Publications/he-who-serves-wins [Accessed March 20, 2016].
37. Porter, M.E., 1996. What is Strategy? *Harvard Business Review*, 74(6), pp.61–78.
38. Keen, P. & Williams, R., 2013. Value Architectures for Digital Business: Beyond the Business Model. *MIS Quarterly*, 37(2), pp.643–647.
39. Zott, C., Amit, R. & Massa, L., 2011. The Business Model: Recent Developments and Future Research. *Journal of Management*, 37(4), pp.1019–1042.
40. Woodard, C.J., Ramasubbu, N., Tschang, F.T. & Sambamurthy, V., 2013. Design Capital and Design Moves: The Logic of Digital Business Strategy. *MIS Q.*, 37(2), pp.537–564.
41. Horlach, B., Drews, P. & Schirmer, I., 2016. Bimodal IT: Business-IT Alignment in the Age of Digital Transformation. In *MKWI 2016 Proceedings*. pp. 1417–1428.
42. Webster, J. & Watson, R.T., 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), pp.xiii–xxiii.
43. Jaubert, M., Marcu, S., Ullrich, M., Malbate, J.-B. & Dela, R., 2014. Going Digital: The Banking Transformation Road Map. *AtKearney*. Available at: http://www.atkearney.de/digital-business/featured-article/-/asset_publisher/featuredarticle/content/going-digital-the-banking-transformation-road-map/10192 [Accessed April 3, 2016].
44. McManus, R. & McDonald, M., 2015. Being Digital: Fast-Forward to the Right Digital Strategy. *Accenture*. Available at: https://www.accenture.com/t20150709T021145__w__/us-en/_acnmedia/Accenture/Conversion-Assets/Blogs/Documents/1/Accenture-Being-Digital-Fast-Forward-2015-Report.pdf [Accessed April 22, 2016].
45. BCG Research, 2015. Digital, for All of Us. *Boston Consulting Group*. Available at: https://www.bcgperspectives.com/content/articles/business-unit-strategy-digital-for-all-of-us-seven-takways-bcg-european-strategy-leadership-summit/ [Accessed April 2, 2016].
46. Riedel, O. & Heinen, M., 2015. Digitalisierung: Wer investiert und profitiert – Wer verliert? *Ernst & Young*. Available at: http://www.ey.com/Publication/vwLUAssets/EY_Studie_-_Digitalisierung_2015/$FILE/EY-Studie-Digitalisierung-2015.pdf [Accessed March 15, 2016].
47. Daub, M. & Wiesinger, A., 2015. Acquiring the Capabilities You Need to Go Digital. *McKinsey*. Available at: http://www.mckinsey.com/insights/business_technology/acquiring_the_capabilities_you_need_to_go_digital [Accessed March 14, 2016].
48. Olanreaju, T., Smaje, K. & Willmott, P., 2014. The Seven Traits of Effective Digital Enterprises. *McKinsey*. Available at: http://www.mckinsey.com/insights/organization/the_seven_traits_of_effective_digital_enterprises [Accessed March 23, 2016].
49. Vollmer, C.A.H., Egol, M. & Sayani, N., 2014. Reimagine Your Enterprise. *Strategy&*. Available at: http://www.strategy-business.com/article/00258?gko=6626b [Accessed April 4, 2016].
50. Basit, T., 2003. Manual or Electronic? The Role of Coding in Qualitative Data Analysis. *Educational Research*, 45(2), pp.143–154.
51. Sproull, N.L., 2002. *Handbook of Research Methods: A Guide for Practitioners and Students in the Social Sciences*, Maryland: Scarecrow Press.
52. Gunasekaran, A., 1999. Agile Manufacturing: A Framework for Research and Development. *International Journal of Production Economics*, 62(1), pp.87–105.
53. Osterwalder, A. & Pigneur, Y., 2010. *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*, New York: John Wiley & Sons, Inc.
54. El Sawy, O.A., Malhotra, A., Park, Y. & Pavlou, P.A., 2010. Seeking the Configurations of Digital Ecodynamics: It Takes Three to Tango. *IS Research*, 21(4), pp.835–848.

The Business Model DNA: Towards an Approach for Predicting Business Model Success

Markus Böhm[1], Jörg Weking[1], Frank Fortunat[1], Simon Müller[2], Isabell Welpe[2], and Helmut Krcmar[1]

[1] Technical University of Munich, Chair for Information Systems, Garching, Germany
{markus.boehm,joerg.weking,frank.fortunat,krcmar}@in.tum.de
[2] Technical University of Munich, Chair for Strategy and Organization, Garching, Germany
{simon.mueller,welpe}@tum.de

Abstract. Business models have gained much interest in the last decade to analyze the potential of new business ventures or possible innovation paths of existing businesses. However, the business model concept has only rarely been used as basis for quantitative empirical studies.
This paper suggests the concept of a Business Model DNA to describe the characteristics of specific business models. This concept allows to analyze business models in order to identify clusters of business models that outperform others and calculate future prospects of specific business models.
We used 181 startups from the USA and Germany and applied data mining techniques, i.e. cluster analysis and Support Vector Machines, to classify different business models in regards to their performance.
Our findings show that 12 distinct business model clusters with different growth expectations and chances of survival exist. We can predict the survival of a venture with an accuracy of 83.6 %.

Keywords: Business Model; Success Prediction; Data Mining; Cluster Analysis; Support Vector Machine

1 Introduction

Business models (BMs) have established in different research communities, like management science and information systems, whereas in practice they are seen as vital for business success [1, 2]. Furthermore, a shift from traditional BMs to electronic ones took place in the last three decades [3, 4]. There are many definitions and frameworks available on the concept today [1].

Business modelling gained significant importance in startup communities, too [2]. Startup firms are a driver of economic growth, innovation and employment opportunities [5]. Unfortunately the failure rate of startups is very high, with estimates ranging from 50% to more than 83% [6, 7]. Why some new ventures fail, while others succeed, is one of the central questions not only for entrepreneurship research, but also for possible entrepreneurs [2, 8]. Scholars agree that current research is still lacking methods to predict firm success [2, 9-11]. Additionally, most of the previous studies

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Böhm, M.; Weking, J.; Fortunat, F.; Müller, S.; Welpe, I.; Krcmar, H. (2017): The Business Model DNA: Towards an Approach for Predicting Business Model Success, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1006-1020

are based on qualitative research [2, 9]. This study proposes a new method for predicting firm success taking into account quantitative measures.

Through the emergence of more powerful data analysis tools and the growing amount of available data, it is possible to use data mining to find meaningful patterns in datasets [12]. Data mining is concerned with making sense of large amounts of data and finding patterns that are difficult to find manually [13].

The goal of this paper is to combine data mining approaches (i.e. cluster analysis and Support Vector Machines (SVM)) with the BM concept to predict the chances of success for startups. This allows better informed, empirically backed strategic and investment decisions [14]. To operationalize this approach, we suggest the BM DNA – in analogy with the human genome – as a concept to describe the characteristics of a specific BM based on various factors. The proposed instance of the BM DNA has been drawn from the 55 BM patterns based on Gassmann, Frankenberger and Csik [15].

The remaining paper is structured as follows. The second chapter describes related work including relevant BM literature. The third chapter elaborates the method that is applied in order to evaluate BMs of startup firms. The fourth chapter illustrates the dataset that has been used. Results of the cluster analysis and the different models of a SVM for classifying BMs are shown in the fifths chapter. The sixth chapter discusses these results and limitations of the study. The final chapter concludes with a short summary, contributions and aspects for future research.

2 Related Work

More than 60 years ago, Drucker [16] defined the term BM as the answer to “who is the customer, what does he value and how can you make money from it”. With the rise of the internet and digital firms the concept has gained more attention in research and practice [1, 10, 17]. This has led to various definitions of BMs in different research streams [18, 19].

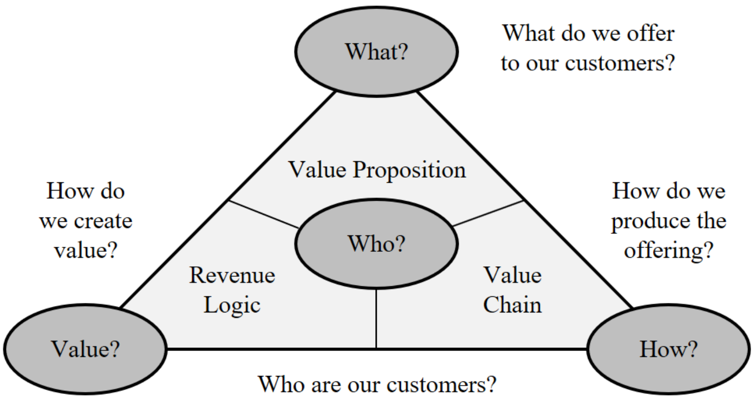


Figure 1. The magic triangle of business models [15]

Drawing from common BM definitions, Gassmann, Frankenberger and Csik [15] have developed a BM framework that consists of four central dimensions: the *Who*, the *Wha*t, the *How*, and the *Value* (see Figure 1). This framework builds the foundation of their 55 BM patterns, which they have identified by analyzing 250 BMs along the four dimensions of the framework. These 55 patterns describe the core of how a BM works. Gassmann, Frankenberger and Csik [15] assume that 90% of the BMs that exist today can be broken down into the same 55 patterns.

Classification of BMs has a long history in entrepreneurship and e-commerce research [20]. Previous studies have used classifications of BMs in empirical research to analyze the influence of BMs on enterprise performance [9]. Lambert and Davidson [9] reviewed research from 1996 to 2010 and found 40 papers that investigate this relationship. Studies that focused on e-commerce conclude that firms should concentrate on interaction platforms for facilitating online transactions and advertising as major revenue stream to be successful. Other studies show that a strategical focus on BM innovation leads to a higher profitability. Additionally, the usage of novelty-centered BMs is related to enterprise success. Moreover, some studies demonstrate that if a BM can easily be transferred to other markets, the firm will more likely be successful [9].

Weill, Malone and Apel [21] distinguish BMs along two dimensions, asset types (financial, physical, intangible, human) and asset rights (creator, distributor, landlord, broker). This leads to a matrix of 16 theoretically possible BM clusters, of which 14 are legal. This framework was used to classify more than 10'000 firms, publically traded on U.S. stock exchanges in order to perform a stock market analysis. They discovered that innovative manufactures are most valued by the stock market. These manufactures are defined as organizations that invest more in research and development than the industry average [21].

Spiegel, Abbassi, Zylka, Schlagwein, Fischbach and Schoder [2] focused on the success of early stage startups using empirical qualitative research. They found that the founders' social capital is crucial to receive funding in this situation. Founders with a better professional social network are able to develop a better BM using their contacts for information and status benefits [2].

However, only few of these studies have focused on quantitative empirical research to quantitatively analyze whether some BMs perform better than others [2, 9]. Scholars agree that future work should aim for the better understanding of relationship of BMs and firm performance [2, 9-11].

3 Research Method

To analyze BM patterns of successful startups, we rely on the Mattermark dataset which is a collection of information about startup firms. In order to avoid a bias towards US or German firms, we drew a random sample of 75 US and 75 German startup firms. For this sample we collected additional information to describe the BM by manually searching the web. Additional information sources included firm websites, press reports and interviews with founders that were publically available.

As the Mattermark dataset is strongly biased towards successful startups we identified additional failed ventures from Crunchbase-Insights, Deadpool and autopsy.io. We sent out a survey to a total of 210 firms and 309 founders asking for additional information to match the data available in the Mattermark dataset. In this way, we were able to include 31 failed startups with sufficient information on financing, revenues, competition and innovativeness of the product.

In the first step of the analysis the BM DNA is built for each of the 181 firms. Similar to the DNA as a molecule that carries most of the genetic information of living organisms, the BM DNA should be a representation of a BMs characteristics. Hence, it should describe the essence of a business in a precise way that allows comparison to other BMs.

In order to build a BM DNA, we draw on the 55 BM patterns developed by Gassmann, Frankenberger and Csik [15]. As 90% of the BMs that exist today can be broken down into the same 55 patterns [15], these BM patterns serve as a sufficient basis for describing the BM DNA. For each of the 55 patterns we used a binary variable to indicate whether a firm uses this pattern or not. In this sense, the BM DNA is a vector that indicates the patterns a certain BM applies. Figure 2 is an exemplary visual representation of the BM DNA. We manually evaluated each firm with regards to the BM patterns it applies. This evaluation and coding process was solely done by one person to ensure a consistent coding of the BM DNAs.

1	2	3	4	5	6		49	50	51	52	53	54	55
0	1	1	0	0	1	...	1	0	1	1	0	0	1

Figure 2. Representation of an exemplary BM DNA

The second step of the data analysis is a cluster analysis based on the BM DNAs. This is not only done to result in meaningful clusters of BMs with different growth perspectives, but also to improve the results of the SVM [22]. We applied the k-means clustering algorithm according to Jain, Murty and Flynn [23] with squared Euclidean distances. The algorithm was used in an iterative process with different numbers of clusters (k) as input factor. The final number of clusters has then been determined with regards to two criteria, the maximum distance between the clusters (I) and the meaningfulness of the clusters (II).

In the third step we developed a metric to evaluate the success of a venture. Therefore, we used both the survival of a firm and its revenue growth. In order to have a comparative metric for revenue growth, we measure the growth relative to the actual revenue (see Figure 4 in the next chapter). The less revenue a firm generates the more revenue growth it needs to generate in order to be evaluated successful.

As a fourth step a SVM is used to classify BMs according to their growth perspectives and whether they are successful or unsuccessful. SVMs use a nonlinear mapping to transform input data (training data) into high dimensional data. The method searches for an optimal, maximum marginal separating hyperplane [24]. This hyperplane is based on support vectors which can be seen as a small subset of the training data [22]. The SVM has been chosen since both neural networks and SVMs have shown to deliver satisfying results in similar studies [25]. An initial comparison of these two techniques using our data indicated better results for SVMs. Thus, SMVs are used in this study with the following information as input:

- BM Information: BM DNA, cluster, scope, focus (B2C or B2B), industry, physical assets, firm age
- Involved people: industry/ foundation experts, investors, founding team size, education of founders, location (Country & City)
- Startup idea: closeness to science and patents, competition and innovativeness

The dataset is randomly split into 80% of the data for training and 20% for testing since this often leads to optimal results [26]. The training data is used to produce the model. This model is then fed with the test data.

The performance of the model is measured by an accuracy, kappa and area under the curve (AUC). The kappa value shows the difference between the calculated solution and a random solution. A kappa value of 0 would indicate a random classification whereas a kappa value of 1 stands for a perfect solution [27]. The AUC measure indicates as well how the model performs in comparison to a random model. The random model would have an AUC of 0.5, whereas an AUC of 1 would indicate a perfect classifier [28].

4 Dataset

The dataset includes 181 firms in total, 31 failed firms resulting from the survey and 150 active firms from the Mattermark dataset. Eighteen of the 31 failed firms were founded in the USA. The databases have been accessed in Mai 2015. The majority of firms is founded between 1999 and 2015. The initial coverage of BM patterns resulting from the manual coding is shown in Figure 3. Almost all firms apply the BM pattern *Digitalization* (11) which stands for digital products or services [15]. However, the dataset is not focused on digital industries.

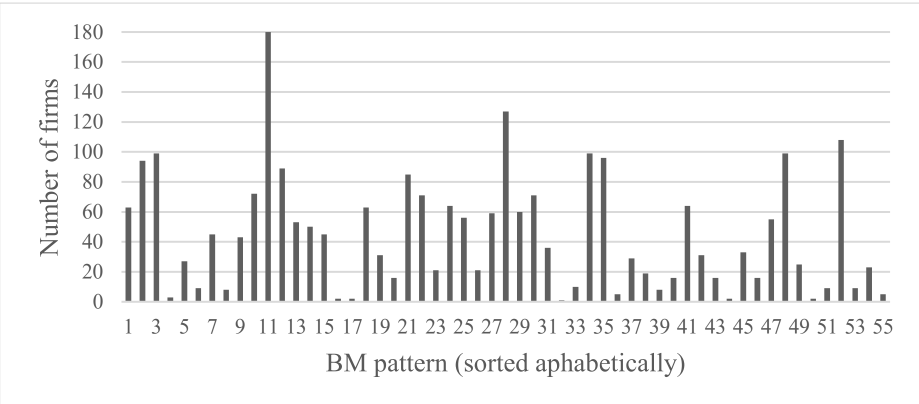


Figure 3. Initial coverage of BM pattern

To get to a meaningful result for the success measure, the data for the USA and Germany had to be treated differently due to the fact that the German firms did not satisfy the same growth criteria than the US firms. This can be linked to the smaller number of average investors and funding the German firms had. Average funding of US startups (238.2 $ M) is more than 3 times higher than the funding German startups received (58.2 $ M). Additionally, German startups had three investors on average while startups from the USA had almost 9 on average.

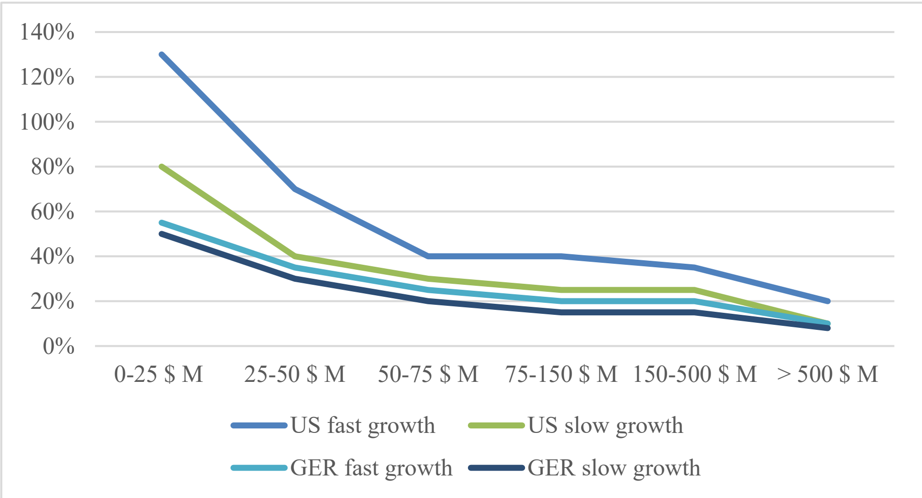


Figure 4. Revenue growth thresholds adapted from [29]

Less funding means fewer possibilities to invest in rapid growth and, therefore, the two regions are not comparable with the same growth percentages. The thresholds were determined in accordance to Maltz and Saljoughian [29] and adjusted for German firms. Figure 4 shows the thresholds that separate the different firms into fast growing and slow growing firms.

5 Predicting Business Model Success

5.1 Business Model Clusters

In the following the characteristics of each of the 12 identified clusters are elaborated. One cluster resembles firms that have a similar combination of BM patterns (BM DNA). Table 1 shows which BM patters are mainly applied by the firms in the respective cluster.

Table 1. Mainly applied business model patterns by cluster

Cluster	*Mainly applied BM patterns*
Freemium Platforms	Freemium and Platforms (Orchestrator, Two-Sided Market or Long Tail)
Experience Crowd Users	Experience Selling, Crowdsourcing and Leverage Customer Data
Long Tail Subscribers	Long Tail and Subscription
Affiliate Markets	Aikido, Affiliation and Platforms (Two-Sided Market, Orchestrator or Long Tail)
Mass Customizing Orchestrators	Mass Customization, Layer Player, Orchestrator and Two-Sided Market
Innovative Platforms	Aikido, Two-Sided Market, Orchestrator and Revenue Sharing
E-Commercer	E-Commerce and Direct Selling
E-Commerce Affiliates	E-Commerce, Affiliate and Long Tail
Add-On Layers	Add-On, Layer Player and Subscription
Crowdsourcing Platforms	Aikido, Crowdsourcing, Customer Loyalty and Platforms (Two-Sided Market or Orchestrator)
Customized Layers	Subscription and Mass Customization
Hidden Revenue Markets	Hidden Revenue, Two-Sided Market, Affiliation and Long Tail

Table 2 indicates the clusters, the number of startup firms in a cluster (n) and their chances of success in terms of survival, fast growth and slow growth. The numbers are to be interpreted as the percentage of firms in the cluster that met the success criteria described in section 3. Thus, firms in the category slow growth also include those firms that managed to grow fast and survival category also includes those firms that managed to grow slowly or fast. The number of members per cluster is well balanced with the fewest members in cluster 0 and 5 with 6. Cluster 7 has the most members with 29.

The *Freemium Platform* cluster mostly includes firms that are using a *Freemium* model or a platform (e.g. *Orchestrator*, *Two-Sided Market* or *Long Tail*). The former offers a free basic service to attract a broad customer base. Revenues are generated by additional chargeable offers. *Orchestrator* coordinate value activities of various firms to offer their customers an aggregated product. The BM of the *Two-Sided Market* serves different customer groups with one platform. The platform is only interesting for any of the two customer groups if the other one is present in a sufficient number as well. For this reason, free services are usually offered to one group of customers. The basis of both BM patterns is a mediation platform. The *Long Tail* pattern is used when the profits for many small payments are achieved by a small margin. Concluding, this cluster includes firms that aggregate the services of different providers and offer a free basic service on their platform as well as a supplementary service. The number of firms in this cluster is rather low (n=6). While all the firms in the sample survive, only half of the firms exhibit strong sales growth. Thus, it is assumed that only a few customers are willing to pay for additional services.

Table 2. Successful business models by cluster

Cluster	*n*	*Fast Growth*	*Slow Growth*	*Survival*
Freemium Platforms	6	50%	50%	100%
Experience Crowd Users	11	18%	27%	64%
Long Tail Subscribers	16	75%	75%	94%
Affiliate Markets	23	52%	65%	78%
Mass Customizing Orchestrators	12	50%	67%	83%
Innovative Platforms	6	67%	67%	100%
E-Commercer	14	43%	57%	57%
E-Commerce Affiliates	29	45%	45%	76%
Add-On Layers	20	45%	60%	90%
Crowdsourcing Platforms	11	73%	91%	91%
Customized Layers	13	54%	62%	69%
Hidden Revenue Markets	20	25%	25%	75%

The *Experience Crowd Users* cluster includes firms that pursue *Experience Selling* and *Crowdsourcing*. In addition, firms in this cluster *Leverage Customer Data*. As part of the *Experience Selling* BM, experience of customers is the key part of the customer value proposition. The goal of the offer is a unique experience. As part of the *Crowdsourcing*, central activities of the value creation are transferred to the crowd, either to the general public or a selected group. This makes it possible to use the potential of large user groups for the firm. The *Leverage Customer Data* pattern is characterized by drawing advantage from customer data. This can be achieved by offering tailored advertising, for example. However, this cluster has the second worst survival rate (64%) and the worst success rate (18%). Our findings suggest that it is very difficult to survive with this BM. Even if startups survive, the growth opportunities are poor. Thus, this cluster is more suitable for niche markets.

The *Long Tail Subscriber* cluster mainly relies on the *Long Tail* and the *Subscription* pattern as an integral part of its strategy. The latter refers to a contractually fixed periodic payment of the user to the provider. In return, the customer can use the offering during a certain period of time. This combination in digital goods is very successful, as duplication of information or provision of software solutions cause almost no additional costs. Therefore, products can be offered very cheap and greater profits can be achieved by many smaller payments over a longer period. Since 94% in this cluster survived and 75% of firms experienced a fast growth, it can be seen as a very successful one.

In addition to the affiliate model, entities within the *Affiliate Markets* cluster utilize the classic patterns of platforms (i.e. *Two-Sided Market*, *Orchestrator* and *Long Tail*). In addition, the *Aikido* pattern is employed. Under the *Affiliation* BM, client referrals to third parties are rewarded with commissions. Price comparison portals are an example of the combination of the affiliate model with a platform. *Aikido* comprises BMs that concentrate on something different than competitors in the same industry. An example is a transfer of BMs that have proven to be successful in other industries. The survival rate of this cluster is average.

The *Mass Customizing Orchestrators* connect the *Mass Customization* pattern with *Layer Player*, the *Orchestrator* and the *Two-Sided Market* pattern. *Mass Customization* means that products, although produced in mass production can be individualized to a certain degree though a variety of options. If a solution for a particular part of the value chain is offered in various industries the *Layer Player* pattern is used. This cluster combines an individual offer with a platform on which the partial offers are available for individual partners. It has a survival rate of 83% and a fast growth rate of 50%. Thus this cluster is quite convenient to grow rapidly.

The next cluster, *Innovative Platforms* focuses besides the classic patterns of a platform such as *Two-Sided Market*, *Orchestrator* and *Revenue Sharing*, mainly on the *Aikido* pattern. Thus, firms of this cluster are trying to integrate new ideas in the BM of platforms. This can be seen as quite successful because all of the firms of the sample have survived until now and two-thirds are also growing fast.

The *E-Commercer* cluster is characterized by the pattern of *E-Commerce* and *Direct Selling*. Entities in *E-Eommerce* offer products or services through the internet. *Direct Selling* means that no intermediaries are used to sell products. This cluster has the lowest survival rate and the second lowest rate of success. This is because the e-commerce market is already saturated. Entering the market is very difficult for young firms without innovative BMs due to strong price competition. An example firm of this cluster is Zalando.

The cluster *E-Commerce Affiliates* performs slightly better than the *E-Commercer* cluster. Besides *E-Commerce* patterns, it follows the *Affiliation* and *Long Tail* pattern. About 3 out of 4 firms survive in this cluster. Only 45% can generate strong growth. By aggregating the offers of many suppliers, a wide selection is created for the customer. Money is generated through links to websites of actual shops.

The *Add-On Layer* uses the *Add-On* pattern, *Layer Player* and *Subscription*. Here, a basic offer is provided relatively cheap. A surcharge must be paid to use more options. This pattern is often used in software as a service (SaaS) products. 90% of firms that use this BM have survived so far. However, only 45% grow fast. This can be attributed to the fact that for many services, the majority of the users just employ basic versions.

Aikido, *Crowdsourcing*, *Customer Loyalty* and the aforementioned platform patterns are used by the cluster *Crowdsourcing Platforms*. *Customer Loyalty* tries to lock-in the customer through incentives such as a bonus scheme or rewards for repeated use. Especially for crowdsourcing, it is extremely important to develop an active community. Therefore, the combination of these two patterns is quite useful. Likewise, a platform for exchange is required for crowdsourcing. The success ensures that users of this BM are right. Nine out of ten firms survive and nearly three out of four grow fast. One potential reason could be that a strong transformation from a pure consumer to dialogue with firms has taken place in recent years. For this reason, many customers are willing to invest time and energy in crowdsourcing campaigns.

The *Customized Layers* also use the *Subscription* and the *Mass Customization* model. The combination of *Layer Player* and *Mass Customization* is very common, especially in the SaaS industry. In this cluster, a hull software is used in many industries but tailored to every firm. We find that slightly less than 75% of the firms survive in this cluster. Additionally, only 54% exhibit strong growth.

The *Hidden Revenue Markets* rely on the *Hidden Revenue* and the *Two-Sided Market* pattern. In addition, *Affiliation* and *Long Tail* are included. In the *Hidden Revenue* pattern there is usually one offer free of charge for one side of a two-sided market. The other group of customers have to pay. The combination of *Long Tail* and *Hidden Revenue* suggests a high volume of transactions. Otherwise it would not be possible to sustain the offer with low payments. A famous example firm of this cluster is the Telegram Messenger. The cluster with this BM does not perform very well. Although three out of four firms survive, only one out of four grows fast.

The 12 clusters show some significant differences. Even though some clusters like the *Freemium Platforms* had a high survival rate, the rate of success was considerably lower compared to other clusters. The same can be seen in regards to the *Innovative Platforms* cluster as well as the *Add-On Layers* cluster. The most consistent cluster was the *Crowdsourcing Platform* cluster which had a survival rate of 91% and a fast growth rate of 73%. This means firms of this cluster that survive tend to grow fast and strong. The *Experience Crowdusers* cluster has a low survival rate of 64% as well as a low success rate of 18%. This cluster can be seen as a niche market. The *E-Commercer* cluster also shows that this is a market with strong competition as most firms do not make it. The few ones that survive grow fast and big. This observation can be seen in the e-commerce market which is dominated by a few big players. The *Long Tail Subscribers* use an interesting BM that combine a big customer base paying a relative low price with a subscription model. In this way the firms are able to attract enough customers to survive and to grow fast. This is as well shown by their survival and success rate of 94% and 75%. Overall it can be seen, that of the 12 clusters 7 are above the 50% survival threshold and only the *Experience Crowdusers* and the *Hidden Revenue Markets* clusters are well below this threshold.

5.2 Success Classification Models

Table 3. Accuracy of success classification

	Model	*calculated*	*true*		*class*	*Kappa*	*Accuracy*	*AUC*
			yes	*no*	*precision*			
1	Fast growth	yes	21	6	77.78%	0.312	66.59%[1]	0.744
		no	12	15	53.57%			
2	Slow growth	yes	26	9	74.29%	0.339	66.8%[1]	0.734
		no	8	12	60.00%			
3	Survival	yes	46	2	95.83%	0.673	83.6%[1]	0.899
		no	2	5	71.43%			
4	Fast growth only BM	yes	17	7	70.83%	0.153	58.4%[1]	0.629
		no	17	14	45.16%			

The SVM is used with different models to classify ventures according to their growth perspectives and survival. Table 3 shows the results of the different classification models. The first model separates firms between fast growing and not fast growing. For example, the model claims that 27 firms are fast growing. However, only 21 of these

[1] Based on 50-50 weighted sample

27 firms are actually classified as fast growing. In other words, their growth relative to their revenue lies above the threshold depicted in Figure 3. The model is able to correctly classify two thirds of the given test data.

The second model takes into account whether a firm is slowly growing or not slowly growing. Hereby, it shows a similar accuracy like the first model. The third model, however, is able to correctly classify 83.6% of given startups in regards to their survival or not. The fourth model is as well concerned with the separation of fast growing firms and not fast growing ones. However, the fourth model uses only information about the BM DNA as an input.

The accuracy considerably varies in the different models. The survival model is the most accurate one as it has a high accuracy as well as a high Kappa measure of 0.673 which is classified as substantial. The AUC measure is also promising with a value of 0.899. The other models lack in accuracy. Furthermore, the success models, models one and two, have Kappa values of around 0.3, which is classified as fair [27]. The AUC values were respectable for both success models though as they were 0.744 and 0.734 respectively.

The fourth model proved that BM information by itself has only very weak classification ability. This means that it is not possible to achieve a high classification accuracy without a combined approach. The low kappa value of 0.153 and AUC value of around 0.6 underlines this statement.

6 Discussion

The paper discusses two data mining approaches for classifying startup BMs in terms of their success by building on a new concept called BM DNA. The first result, arising from a cluster analysis, shows currently promising or not promising BM clusters. As part of the second result a SVM is introduced for classifying BMs as successful or unsuccessful.

The k-means clustering algorithm is used in this study since it has shown to be very efficient [23]. However, the algorithm shows some downsides. The number of clusters is an input factor, the outcome is depended on the initial solution, it is sensitive to outliers and can end up in local optimal solutions instead of global ones [23, 30]. In addition, the Euclidian distance used with binary data can be seen as problematic [23]. The iterative usage mitigates these disadvantages. Alternatively, a hieratical agglomerative clustering algorithm can be used and is seen as future research.

The clusters achieved a satisfactory result even though some similar patterns correlate with different clusters. These patterns are applied in many companies. On the other hand, some patterns, like *Cash Machine*, *Trash-to-Cash*, *Target the Poor*, *White Label* or *Ultimate Luxury*, do not show a strong correlation with any pattern. There were not enough firms that use these patterns in the sample.

There are different models in the literature aiming to classify BMs in regards to survival and not survival (Table 3). The survival model (3) correctly predicts 95.83% of the successful firms and 71.43% of the unsuccessful firms without considering economic anomalies. This accuracy is as good as or even better than most studies of

Business Failure Prediction (BFP) for new ventures. Lussier [31] was able to match 73% of the failed and 65% of the successful firms by using logistic regression.

This study used SVM as classification method. Different methods are used in literature. Gartner, Starr and Bhat [32] used multiple discriminant analysis and were able to correctly classify 85% of the firms which is similar to this approach. Marom and Lussier [33] achieved an accuracy of 85.4% which is also comparable. However, our model is superior in classifying successful businesses as Marom and Lussier [33] were only able to classify the successful ones with an accuracy of 84%. Ciampi and Gordini [34] were able to correctly classify 68.4% of their firms. Ecer achieved results of 91.18% and 88.24% for neural networks and SVMs respectively. Hence, both techniques have a high prediction ability [25]. Ecer [25] achieved a better accuracy using neural networks while Olson, Delen and Meng [24] had the best result with decision trees. Wilson and Sharda [35] used the same financial ratios as Altman since they wanted to compare neural networks and multiple discriminant analysis. Here, multiple discriminant analysis was outperformed by neural networks in every test. However, in this study, the prediction ability of the SVM proved to be superior for the survival model in comparison to other methods in literature.

There are some limitations of SVMs. Information about the founders and their prior knowledge of the industry and founding, the degree of innovativeness of the new firm as well as competition, founding team size, and patents improved the model accuracy [7, 31, 36-38]. While the variables did produce a satisfactory model, the problem of transparency and transportability, the inherent issue of all SVMs, still exists [24]. With SVMs, it is not possible to show how much each variable is influencing the outcome in a numerical way like with a logistic regression. Unfortunately, it is also not possible to visualize the model since it has too many dimensions for producing a graphical representation. Other authors used techniques like ANOVA, principal component analysis (PCA) or correlation matrices to preselect the variables they used in their models [39, 40]. However, these approaches neglect the interaction of the independent variables among themselves and can lead to a worse selection of variables. Thus, SVM has been used in this study and delivers very good results compared to other studies.

While the survival model was able to deliver a good prediction accuracy, models used on mature firms that include profound financial ratios are superior in accuracy. The model by Altman was able to correctly classify 95% of the cases. Edminster also used a number of financial ratios in his multiple discriminant analysis model and had a 92% classification success rate on small firms [41].

However, the focus of this paper is on startup firms. The rate of failure in startups is considerably higher than in mature firms [2]. Additionally, in early stage startups, it might be difficult to get profound financial numbers. Thus, a success model with an accuracy of almost 85% is a very good classification model for startups.

The proposed model in this paper is a considerable contribution for startup classification and success prediction. However, this study is a proof of concept only. Our study showed that the used method is able to correctly classify startups concerning their overall success. More data is needed to clearly verify, enhance and refine the model.

7 Conclusion

This paper suggests a BM DNA as means to describe the specific characteristics of certain BMs. It demonstrated the applicability of this concept by drawing from the 55 BM patterns based on Gassmann, Frankenberger and Csik [15] to describe the characteristics of a sample of 181 firms.

The BM DNA was then used in a cluster analysis. 12 different BM cluster were created including their chances for survival, for slow growth and for fast growth. These 12 cluster can be used in practice to make an assumption for growth perspectives of an actual startup. Some BM clusters seem to have better future prospects than others.

In a next step a SVM has been used to estimate a BMs survival and growth perspectives. The classification of survival worked fairly well. It is a major contribution for research since comparable studies do not achieve similar accuracies. Furthermore, an advanced model should be used in practice by entrepreneurs and investors as a step of business model evaluation.

However, the classification of slow and fast growth as a means to evaluate the success of a BM was not very reliable. Future research should investigate in additional classification algorithms (1), apply the model on a larger dataset (2) and extend the BM DNA with additional information (3). Our fourth model shows that this is needed to correctly classify BMs concerning their growth perspectives.

While future versions of the BM DNA serve as a good predictor of BM success, it is still a reflection of data from the past and, thus, past success factors. Innovation and consequently the chances for exceptional success will remain a creative process of identifying new approaches and promising re-combinations. However, predictions based on the BM DNA may serve as a sounding board for entrepreneurs and investors to critically reflect on specific BMs and make purposeful decisions.

8 Acknowledgements

This paper originated in the context of the study "Business Models in the Digital Economy – A Study of the German Innovation System" from the Commission of Experts for Research and Innovation [42]. Parts of the paper at hand have been used in this study.

References

1. Wirtz, B.W., Pistoia, A., Ullrich, S., Göttel, V.: Business Models: Origin, Development and Future Research Perspectives. Long Range Planning 49, 36-54 (2016)
2. Spiegel, O., Abbassi, P., Zylka, M.P., Schlagwein, D., Fischbach, K., Schoder, D.: Business Model Development, Founders' Social Capital and the Success of Early Stage Internet Start-Ups: a Mixed-Method Study. Information Systems Journal 26, 421-449 (2015)
3. Baden-Fuller, C., Morgan, M.S.: Business Models as Models. Long Range Planning 43, 156-171 (2010)

4. Magretta, J.: Why Business Models Matter. Harvard Business Business Review 80, 86-92 (2002)
5. Yankov, B.: Overview of Success Prediction Models for New Ventures. International Conference on Automatics and Informatics 12, 13-16 (2012)
6. Laitinen, E.K.: Prediction of Failure of a Newly Founded Firm. Journal of Business Venturing 7, 323-340 (1992)
7. Wetter, E., Wennberg, K.: Improving Business Failure Prediction for New Firms: Benchmarking Financial Models with Human and Social Capital. Journal of Private Equity 12, 30-37 (2009)
8. Cooper, A.C.: Challenges in Predicting New Firm Performance. Journal of Business Venturing 8, 241-253 (1993)
9. Lambert, S.C., Davidson, R.A.: Applications of the Business Model in Studies of Enterprise Success, Innovation and Classification: An Analysis of Empirical Research from 1996 to 2010. European Management Journal 31, 668-681 (2013)
10. Al-Debei, M.M., Avison, D.: Developing a Unified Framework of the Business Model Concept. European Journal of Information Systems 19, 359-376 (2010)
11. George, G., Bock, A.J.: The Business Model in Practice and its Implications for Entrepreneurship Research. Entrepreneurship Theory and Practice 35, 83-111 (2011)
12. Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., Tufano, P.: Analytics: The Real World Use of Big Data. IBM Institute for Business Value (2012)
13. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From Data Mining to Knowledge Discovery in Databases. AI Magazine 17, 37 (1996)
14. McAfee, A., Brynjolfsson, E.: Big Data: The Management Revolution. (cover story). Harvard Business Review 90, 60-68 (2012)
15. Gassmann, O., Frankenberger, K., Csik, M.: The Business Model Navigator: 55 Models That Will Revolutionise Your Business. Pearson, Harlow (2014)
16. Drucker, P.: The Practice of Management. Harper & Row, New York (1954)
17. Chesbrough, H., Rosenbloom, R.S.: The Role of the Business Model in Capturing Value from Innovation: Evidence from Xerox Corporation's Technology Spin-Off Companies. Industrial and Corporate Change 11, 529-555 (2002)
18. Krcmar, H., Böhm, M., Friesike, S., Schildhauer, T.: Innovation, Society and Business : Internet-based Business Models and their Implications. 1st Berlin Symposium on Internet and Society, 1-33 (2011)
19. Zott, C., Amit, R., Massa, L.: The Business Model: Recent Developments and Future Research. J Manage 37, 1019-1042 (2011)
20. Weiner, N., Renner, T., Kett, H.: Geschäftsmodelle im Internet der Dienste: Aktueller Stand in Forschung und Praxis. Fraunhofer Verlag, Stuttgart (2010)
21. Weill, P., Malone, T.W., Apel, T.G.: The Business Models Investors Prefer. Mit Sloan Manage Rev 52, 16-20 (2011)
22. Çomak, E., Arslan, A.: A new Training Method for Support Vector Machines: Clustering k-NN Support Vector Machines. Expert Systems with Applications 35, 564-568 (2008)
23. Jain, A.K., Murty, M.N., Flynn, P.J.: Data Clustering: A Review. ACM Computing Surveys 31, 264-323 (1999)
24. Olson, D.L., Delen, D., Meng, Y.: Comparative Analysis of Data Mining Methods for Bankruptcy Prediction. Decision Support Systems 52, 464-473 (2012)

25. Ecer, F.: Comparing the Bank Failure Prediction Performance of Neural Networks and Support Vector Machines: The Turkish Case. Ekonomska istraživanja 26, 81-98 (2013)
26. Wang, J., Neskovic, P., Cooper, L.N.: Training Data Selection for Support Vector Machines. International Conference on Natural Computation, 554-564. Springer (2005)
27. Viera, A.J., Garrett, J.M.: Understanding Interobserver Agreement: The Kappa Statistic. Family Medicine 37, 360-363 (2005)
28. Pruessner, J.C., Kirschbaum, C., Meinlschmid, G., Hellhammer, D.H.: Two Formulas For Computation of the Area Under the Curve Represent Measures of Total Hormone Concentration Versus Time-Dependent Change. Psychoneuroendocrinology 28, 916-931 (2003)
29. Techcrunch, How Fast Should You Be Growing?, http://techcrunch.com/2013/08/24/how-fast-should-you-be-growing (Accessed: 21.07.2016)
30. Rai, P., Singh, S.: A Survey of Clustering Techniques. International Journal of Computer Applications 7, 1-5 (2010)
31. Lussier, R.N.: A Nonfinancial Business Success versus Failure Prediction Model for Young Firms. Journal of Small Business Management 33, 8-20 (1995)
32. Gartner, W.B., Starr, J.A., Bhat, S.: Predicting New Venture Survival: An Analysis of "Anatomy of a Start-Up." Cases from Inc. Magazine. Journal of Business Venturing 14, 215-232 (1999)
33. Marom, S., Lussier, R.N.: A Business Success Versus Failure Prediction Model for Small Businesses in Israel. Business and Economic Research 4, 63-81 (2014)
34. Ciampi, F., Gordini, N.: Small Enterprise Default Prediction Modeling through Artificial Neural Networks: An Empirical Analysis of Italian Small Enterprises. Journal of Small Business Management 51, 23-45 (2013)
35. Wilson, R.L., Sharda, R.: Bankruptcy Prediction Using Neural Networks. Decision Support Systems 11, 545-557 (1994)
36. Hyytinen, A., Pajarinen, M., Rouvinen, P.: Does Innovativeness Reduce Startup Survival Rates? Journal of Business Venturing 30, 564-581 (2015)
37. Song, M., Podoynitsyna, K., van der Bij, H., Halman, J.I.M.: Success Factors in New Ventures: A Meta-analysis. Journal of Product Innovation Management 25, 7-27 (2008)
38. Cassar, G.: Industry and Startup Experience on Entrepreneur Forecast Performance in new Firms. Journal of Business Venturing 29, 137-151 (2014)
39. Lee, K.C., Han, I., Kwon, Y.: Hybrid Neural Network Models for Bankruptcy Predictions. Decision Support Systems 18, 63-72 (1996)
40. Dimitras, A.I., Slowinski, R., Susmaga, R., Zopounidis, C.: Business Failure Prediction Using Rough Sets. European Journal of Operational Research 114, 263-280 (1999)
41. Edmister, R.O.: An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. Journal of Financial and Quantitative Analysis 7, 1477-1493 (1972)
42. Müller, S., Böhm, M., Schröer, M., Bakhirev, A., Baiasu, B., Krcmar, H., Welpe, I.: Geschäftsmodelle in der Digitalen Wirtschaft, Studien zum Deutschen Innovationssystem. Berlin: EFI (2016)

“I did use it!” - Assessing subjective vs objective cognitive artifact usage

Stefan Morana[1], Julia Kroenung[2], and Alexander Maedche[1]

[1] Karlsruhe Institute of Technology, Institute of Information Systems and Marketing (IISM) and Karlsruhe Service Research Institute (KSRI), Karlsruhe, Germany
{stefan.morana,alexander.maedche}@kit.edu

[2] University of Mannheim, Dieter Schwarz Foundation Endowed Chair for E-Business and E-Government, Mannheim, Germany
{kroenung}@bwl.uni-mannheim.de

Abstract. For decades, IS researchers have discussed the reliability of subjective measurements to assess actual artifact usage. Especially in experimental settings, as in the design science context for instance, the participants’ usage data of the evaluated artifact represents an important point of analysis. However, collecting objective usage data, (i.e. logfiles) is often not feasible depending on the artifact. In this paper, we present the theoretical grounding of collecting cognitive artifact usage data using eye-tracking technology. Grounded in immediacy and eye-mind assumption the participants’ artifact fixations are used as objective usage measurement. The question remains if in comparison, the collection of subjective (e.g. perceptual) usage data is sufficient and reliable for such experiments. The results of our comparative analysis indicate that researchers could use subjective measurements when comparing different artifact designs and should rely on objective measurements when testing the effect of an artifact compared to a control group without artifacts.

Keywords: Artifact, eye-tracking technology, laboratory experiments, objective usage, subjective usage.

1 Introduction

The use of information technology (IT) by individuals remains the variable of interest in the information systems (IS) field and has been recognized as a key element of the *missing link* from IT investments to business value [1, 2]. The question if a certain IT artifact is used or not and which outcomes are related to the IT usage, has been subject of investigation in numerous experiments in IS research. In general, there are two possible approaches for measuring usage: subjective or objective measurements [3]. Already in 1989, Fred Davis concluded that “n*ot enough is currently known about how accurately self-reports reflect actual behavior*” [4: p. 334] and still, more than 25 years later, there is an ongoing debate on the reliability of self-reported, subjective usage measurements.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Morana, S.; Kroenung, J.; Maedche, A. (2017): "I did use it!" - Assessing subjective vs objective cognitive artifact usage, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1021-1035

Overall, there is a strong tendency that subjective data is no fully reliable approximation of objective data in the research discourse [5]. Still, there are inconsistent results that argue for the reliability of subjective usage [6, 7].

The primary objective of experiments in the design science research (DSR) [8] context is the investigation if certain effects occur while participants use an IT artifact (we use the term artifact to subsume the terms artifact / treatment / intervention) [9]. In DSR, the usage of the evaluated artifact in the experiment is an important information that requires special attention [10]. Within an experimental setting, researchers can collect subjective data (e.g. by asking the participants on their artifact's usage in a questionnaire) or objective data (e.g. by recording the participants' interaction with the artifact using a screen cam). In most experimental settings, the participants actively use the artifact, from a human-computer interaction (HCI) point of view, for example, by clicking on a button or scrolling on a website. In such experimental settings, assessing the participants' objective usage is easy. For an overview on objective use measurements in experiments see Eckhardt et al. [3].

In contrast to this, there are experimental settings in which the participants do not actively use the artifact (e.g. clicking a button) during the experiment, but use it from a cognitive point of view. One example is the provision of a textual explanation [11] (as the actual artifact) in an experimental application. The participants can read these explanations during the experiment and use the provided textual information without any active interactions with the computer. Rather, the participants read the information, store it in their memory, and use it, for example, to process the experimental task. We define this as cognitive artifact usage; the user interacts with the artifact only by viewing at it and processing the provided information, for example, text or images, cognitively. Due to missing direct human-computer interactions with the artifact (e.g. by counting click rates), the logging of such objective usage data is more difficult. However, in most cases, it is important to know if and to which extent the participants used the artifact during the experiment to assess its effect or impact. In such experimental settings, researchers often rely on subjective usage data, because objectively measuring such cognitive artifact usage is difficult.

A possibility to address this shortcoming is the utilization of eye-tracking technology. Eye-tracking technology enables researchers to measure the participants' eye movements on the screen and thus, allow an objective measurement of the users' cognitive artifact usage. According to the immediacy assumption and the eye-mind assumption [12], researchers can infer the participants' current cognitive activity based on their current glance on the computer. The following example shall illustrate this assumption: While looking at a picture in a gallery, the individual most likely also thinks about the picture and thus, cognitive processes take place. Consequently, eye-tracking technology enables researchers to derive the participants' current cognitive activity by recording their eye fixations during the experiment. However, the utilization of eye-tracking technology in experiments today is still rather expensive from both, a financial and an effort point of view, compared to gathering subjective data using a survey. Thus, gathering subjective usage data is less effort. Considering the ongoing debate on subjective vs. objective measurement, the question remains if subjective measurement data is reliable for assessing cognitive artifact usage in laboratory experiments as formulated:

Do subjective measurements reliably approximate objective measurements with respect to cognitive artifact usage in laboratory experiments?

Our research contributes to the ongoing debate if subjective usage measurements are a reliable approximation for objective usage measurements in laboratory experiments. We provide insights from our DSR laboratory experiment and discuss under which conditions the application of subjective measurements is sufficient and when objective measurements are required based on our experiment data. From a methodological point of view, we present the theoretical baseline for objectively measuring cognitive artifact usage in laboratory experiments with eye-tracking technology, grounded in the immediacy assumption and the eye-mind assumption [12]. We demonstrate how to apply eye-tracking technology for measuring and analyzing objective artifact usage. In addition, our findings contribute to the IT usage research stream as they provide an example under which conditions subjective usage measurements can be applied.

2 Theoretical Foundations and Related Work

2.1 Information Systems

Researchers conduct experiments to evaluate if the usage of a certain artifact results in the proposed effect or not. There are two types of usage measurements within experimental setting, participants report their subjective usage, or researchers measure the usage objectively during the experiment. Wu and Du [13], based on a comprehensive literature review on IS usage research, discuss two different types of subjective usage types: reported usage and assessed usage. The researchers argue that the key difference between reported usage and assessed usage is the application of ordinal scales for the measurement. Reported usage indicates the "*users-reported amount of time or number of times of using a system*." [13: p. 683] In contrast to this, assessed usage "*refers to the ordinal-scale-measured intensity and extent of using a system*." [13: p. 683] For the remainder of the paper, we refer to subjective usage for both type of usage types.

Using subjective (reported and assessed) usage measurements can result in non-reliable data and distorted results, due to possible common method biases [7, 14]. There is an ongoing debate whether subjective usage is a reliable approximation for objective usage. When talking about artifact usage in experiments, the whole stream on IS usage research is important to consider. Before summarizing a selection of findings on the subjective vs. objective usage debate, we shortly address the construct system usage itself. In decades of IS usage research there is few "*in-depth, theoretical assessment of the construct*" [15: p. 228], until it was re-conceptualized by Burton-Jones and Straub [15]. The researchers propose six measures for system usage ranging from very lean to very rich, enabling researchers to provide more insights in the applied usage construct. Straub et al. [7] also discuss the assessment of system usage and the comparability of subjective and objective system usage. They found, in contrast to their expectations, that both types of usage measurements are not strongly correlated to each other [7].

In a repetition of the study by Straub et al. [7], Barnett et al. [6] found a stronger relationship between the subjective and objective system usage. Their analysis reveals a correlation between the perceived and objective usage, ranging from $r = 0.38$ to 0.43 [6]. The researchers conclude their article with a suggestion to consider the context of the study when deciding which type of usage measurements to apply. Barnett et al. [6] suggest to apply objective data when the actual usage is connected to the dependent variable (such as performance) and subjective data when general insights are required [6: p. 81]. In the discussion of their empirical results, Burton-Jones and Straub [15] also shortly address the question when to use subjective and objective measures. They suggest objective usage "*to measure the system and task aspects of usage*" and subjective usage (such as questionnaires) "*to measure user states such as cognitions or emotions during usage*". [15: p. 243]

2.2 Psychology

Investigating the discrepancy between self-reported and actual behavior has a tradition in experimental psychology and is related to the debate on usage behavior in IS research. There is evidence that participants are unaware of many biases in their choices resulting in questioning the reliability of subjective behavior compared to objective behavior [16]. One stream of research in this discipline further addresses introspective information and the reliability of such subjective data for decades. The reliability of subjective data measurements depends on the applied methods. Verbal reports for instance, can be accurate and a significant reflection of the objective behavior, if the introspective information is relevant to the current task and available in the participants' short-term memory [17, 18]. Thus, introspection can only access cognitive representation kept in the participants' working memory [17]. Other information influences the participants' introspective information. Studies revealed, the participants subjective measurements can be influenced by the perceived difficulty of the task [19] or subjects' own motor responses [20].

2.3 Human-Computer Interaction

In contrast to these supportive findings on a correlation between subjective and objective usage data in psychology, there are also non-supportive results. Analyzing participants' mobile application usage, Reuver and Bouwman [5] found no correlation between subjective and objective usage measurements. Moreover, they discuss the possibility of type I and type II errors when relying on subjective measurements alone [5]. All in common, in the aforementioned usage studies is the participants' active interaction with the system, for example, Straub et al. [7] investigate voice mail usage and Reuver and Bouwman [5] investigate mobile applications use. In such settings, the objective artifact's usage measurement is possible, for example, by using logging functionality within the applications.

This is not possible for assessing the artifacts' usage without some sort of direct human-computer interaction, for example, clicking on a certain button. An example of such a setting is the provision of an explanation within the experiment's application, which is always present and there is no activity by the participants required to make the explanation visible. The participant can read the explanation or not, there is no possibility to log this reading process within the experiment's application. The participant is looking at the explanation, reading it, and cognitively processing the provided information. Eye-tracking technology enables researchers to capture this reading or looking process by recording the participants' eye movements on the screen [21, 22]. This is grounded in the immediacy assumption and the eye-mind assumption proposed by Just and Carpenter [12, 22]. The immediacy assumption "*is that a reader tries to interpret each content word of a text as it is encountered.*" [12: p. 330] The eye-mind assumption "*is that the eye remains fixated on a word as long as the word is being processed.*" [12: p. 330]

Eye-tracking technology is also used for quite some years in HCI research [23] and psychological research [21, 24–26]. Several studies showed the linkage between the humans' eye movements and their attention [22, 24]. With respect to this research, we focus on this very link between eye fixations and attention, since the link between fixations and cognitive processing lies beyond the scope of this article. Although research in the field of cognitive psychology regards eye-fixations also as approximation for cognitive processing [24]. Eye-tracking technology is, for example, used to investigate eye movements in reading and information processing research [24, 25]. In a recent study, the participants' subjective report of eye fixations was compared to the objectively measured fixations using eye-tracking technology [27]. In this study, the participants were able to report at least a subset of their eye movements. Nevertheless, the researchers also identified false reported eye movements [27]. In line with the results on subjective vs. objective usage measurements [6], there is the potential for subjective measurements to be a reliable approximation for objective measurements, but it depends on the study context and applied methods. The eye-tracking technology provides researchers "*quantitative and qualitative measures of observers' subjective reports and reveal experimental effects of visual search that would otherwise be inaccessible." [27: p. 1]* Depending on the experimental setting, participants might not be able to provide a justification for their own decisions [28] and thus, eye-tracking technology can be used to obtain objective data [21].

3 Hypothesis development

Summarizing the brief presentation of the ongoing discussion on subjective vs. objective usage measurement, there are findings supporting a correlation [6, 7] and findings rejecting a correlation [5]. In line with the supportive findings, we assume that a reliable rating of the participants' cognitive artifact usage in a laboratory experiment is possible. In such a setting, the participants use the artifact cognitively, for example, by reading or viewing at it, based on the eye-mind assumption [12]. We assume these cognitive processes are stored within their short-term memory. According to the immediacy assumption [12], participants should be able to recall their recently conducted activities. Thus, when asking the participants immediately after the experiment, their activities during the course of the experiment should be available in their short-term memory. This relationship is additionally grounded in research on introspective information [17, 18]. Thus, we argue the participants' perception of their artifact usage should be possible. We propose that the subjective artifact usage is a reliable measurement in experimental settings with cognitive artifact use. Summarizing our assumptions, we propose the following hypothesis:

Hypothesis*: Subjective measurements are positively related to objective measurements of cognitive artifact usage in laboratory experiments.*

For testing this hypothesis, we consider the objective artifact usage as the reliable measurement. We test if there is a significant positive relationship between the subjective perceptions of artifact usage and its objective measurements. This would indicate that the perceived usage measurements are a reliable approximation for the objective usage measurements as tested in several other studies [5–7]. In order to prevent possible type I and II errors [5], we also test for a relationship between the perceived artifact's usage and the objective usage of other applications used in the experiment. This test is required in order to ensure that the participants' subjective measurement, their introspective information, only reflects the artifact's usage and not the usage of other applications. Next, we describe the experimental setup for testing our hypothesis.

4 Experimental setup

We test our hypothesis in the context of an experiment on the effect of providing textual and visual explanations [11] to support individuals' process execution. The experiment itself is part of a DSR project aiming to design an assistance system [29] to support users' process knowledge and process execution performance [30]. In the experiment, we simulated an IT ticketing process and tested if the explanations' provision by the assistance system has the intended effects on the users. The participants had to handle eight IT requests (e.g. create a new user account or purchase of hardware) provided by a simulated email client in a simulated ticketing system according to the specified ticketing process. This ticketing process is a simplified version of a real-world organizational ticketing process adapted from the DSR project's case company. The applications used in the experiment (see **Figure 2**) are simplified versions of the case company's applications used to handle their ticketing processes. One week before the experiment, the participants received a training on the experimental ticketing process and applications. They also received the documentation and we asked them to study it in detail.

The participants in the experiment are students (bachelor and master level) from a large German university and we randomly assigned them to one of three groups. We developed two versions of the assistance system as treatments. In both versions, the assistance system is depicting the sequence of the ticketing process steps supporting the participants with its execution. Moreover, the "advanced treatment" additionally provides textual explanations [11] for each of the process steps. In the following, the group “advanced treatment” refers to the group receiving the advanced assistance system, the group “basic treatment” refers to the group receiving the basic version of the assistance system (without the additional textual explanations). We decided to test two assistant systems with varying amount of text to control for possible differences in the participants usage behavior. The control group received an empty assistance system only stating that the participants should process the emails according to the process specifications. The usage of the assistance systems was voluntary for the participants, but the provided information supported the execution of the experiment’s tasks. We refer to the email client, ticket system, and service catalog as the “experiment’s applications” in the reminder of the paper.

To anonymize the paper, we renamed the assistance system to “ARTIFACTNAME”. In the experiment, the training upfront, and in the survey afterwards, the assistance system had a dedicated name enabling the participants to identify, distinguish, and remember it when reporting their subjective usage. We measured the subjective usage perceptions by two self-developed questions on 7-point Likert scale (intervals in arrow brackets), adapted from the subjective usage measures applied by Davis [4]:

Question use 1: *I used the* ARTIFACTNAME *in the experiment*
<strongly disagree – strongly agree>

Question use 2: *In the experiment, I used the* ARTIFACTNAME
<never – very frequently>

We used Tobii X2-30 eye tracker and the software Tobii Studio for recording the experiment’s sessions. Eye-trackers by Tobii are used in many research domains and Tobii is one of the market leaders of eye-tracking technology world-wide [31]. The eye-tracker recorded the participants’ eye movements with a frequency of 30 Hz. Thus, every 33ms the participants’ eye fixation and position on the screen, for example, was measured and stored with the timestamp by Tobii Studio. In addition, the experiment’s sessions were screen-recorded. The resulting videos as well as the collected data was analyzed subsequently. Following Figure 1 depicts a screenshots of the recorded sessions and the highlighted eye movements of the participants (red dots and lines) based on the collected data by Tobii Studio. The diameter of the red dot indicates the duration of the participants’ eye fixation at the certain point on the screen and the red lines indicate the movement of the participants’ eyes on the screen. Such a visualization of the participants’ eye movements and fixation durations supports researchers to reconstruct the participants’ behavior during the experiment. The screenshot shows a participant with the “advanced treatment” providing the sequence of the ticketing process and the additional explanations for each process step. Please note, to anonymize this paper, the artifact name (top right) and the image of the participant (top left) was blurred.

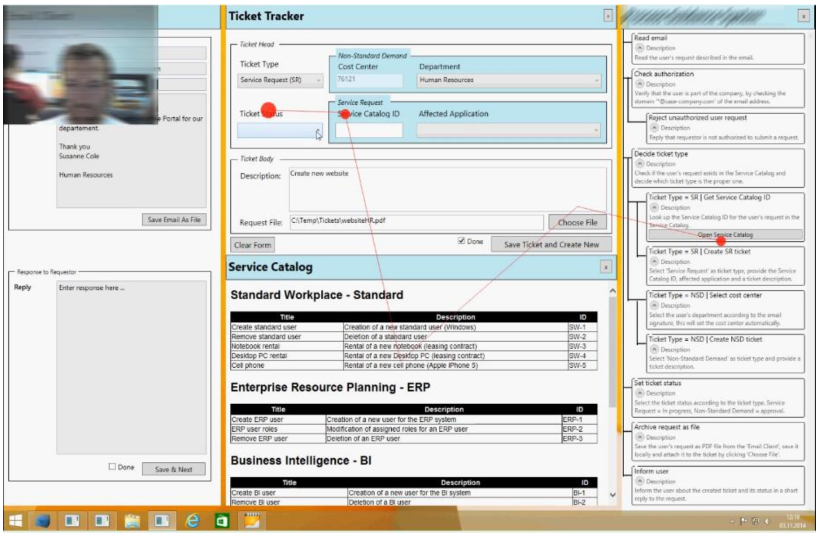


Figure 1: Screenshots of experiment recordings

Within Tobii Studio, we defined four areas of interest (AoI), for the experimental applications (see **Figure 2**). The software analyzes the recordings and provides the participants' fixation counts for each AoI. We used these fixation counts as objective usage measurements grounded by the immediacy assumption and the eye-mind assumption [12].

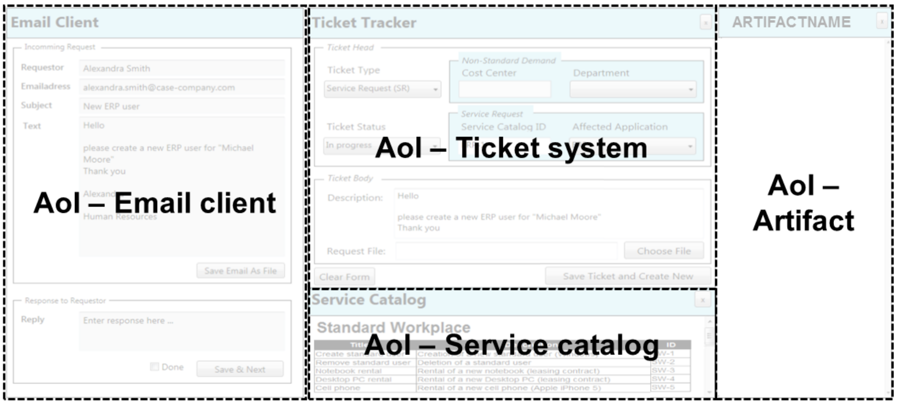


Figure 2: Experimental applications and AoIs

Following Burton-Jones and Straub [15], both usage measurements address the "extent of use" and refer to the second type of measurement richness [15]. Please note we do not discuss the experiment's actual measurements and results here, as we focus on the question whether subjective usage measurements are a reliable approximation of objective usage measurements. However, during the experiment we controlled for the intended effect of the assistance systems and the results indicate a positive effect of both assistance systems compared to the control group. Next, we present the analysis of the participants' subjective and objective usage in the experiment.

5 Analysis

In total, 118 students participated in the experiment, 29 females and 89 males with an average age of 21.81 years (standard deviation = 2.31 years). Table 1 contains the mean values and standard deviations of the total fixations, time, and the survey data for the three experimental groups.

We applied a Shapiro-Wilk test [32] to test if our data follow a normal distribution. All measurements, except for three measures, are not normally distributed, thus referred to non-parametric statistical tests for the data analysis. The mean values of the total fixations indicate a similar usage of the email client, ticket system, and service catalog among the three groups. The artifact's usage varied strongly between the three groups. The highest mean usage, both subjective and objective, has the "advanced treatment" group with a mean count of 2417.97 fixations that equals to 80.59s compared to a mean fixation count of 356.80 fixations or 11.89s of the control group.

Table 1. Mean values and standard deviation of the total fixations, time, and survey data

			Objective measurement					Subjective measurement	
Group	n		AoI – Artifact[1]	AoI - Email client[1]	AoI - Ticket system[1]	AoI - Service catalog[1]	Time[2]	Survey - Use 1[3]	Survey - Use 2[3]
Advanced treatment	39	mean	2417.97	6221.92	5634.18	2520.97	715.90	4.41	4.03
		SD	3141.65	1991.18	1905.30	1107.07	288.61	1.85	1.91
Basic treatment	38	mean	1148.60	6357.26	6176.66	2637.37	753.41	3.97	3.50
		SD	1071.33	2598.87	2363.80	1704.95	307.86	2.11	2.01
Control	41	mean	356.80	6899.10	6115.24	2238.24	684.96	2.76	2.54
		SD	297.07	2101.68	2188.53	1312.69	195.76	1.98	1.91
[1] measured in fixations | [2] measured in seconds | [3] measured on a 7-point Likert scale from 1 to 7									

As expected, the treatment groups' usage is higher than the control group's usage. In order to gain deeper insights, we calculated the fixations per minutes for the four AoI and visualized them in boxplots (see Figure 3). The results indicate a similar usage of the ticket system and the service catalog among the three groups. Surprisingly, we found a significant difference for the email client usage (measured in fixations per minute) between the "advanced treatment" group and the control group (p-value = 0.016) as well as between the "basic treatment" group and the control group (p-value = 0.012).

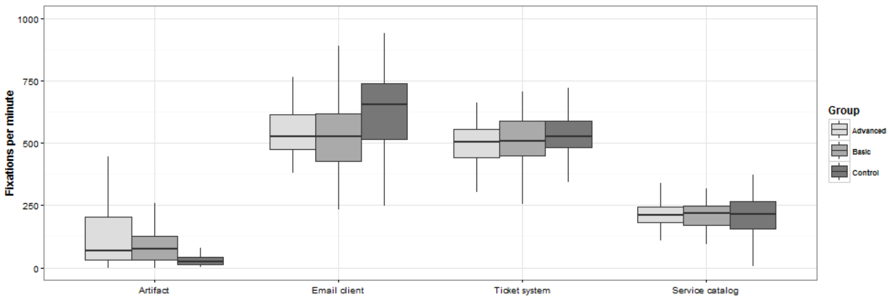


Figure 3. Participants' fixations per minute

The analysis of the artifact's fixations per minute reveals a similar result. For both treatment groups there is a significant difference compared to the control group (both p-values < 0.001). Next, we test for the positive correlation between the objective usage and the subjective usage with the Spearman Rho test [33, 34]. Table 2 contains the Rho values of the correlation tests and the correlations' significance.

Table 2. Correlations between objective and subjective usage (means)

	Complete dataset				**Advanced treatment group**			
	AoI-A	AoI-E	AoI-T	AoI-S	AoI-A	AoI-E	AoI-T	AoI-S
AoI Artifact	1.00 ***	0.46 ***	0.59 ***	0.30 **	1.00 ***	0.66 ***	0.78 ***	0.43 **
Survey - Use 1	0.52 ***	0.24 **	0.28 **	0.19 *	0.68 ***	0.49 **	0.62 ***	0.44 **
Survey - Use 2	0.48 ***	0.24 **	0.26 **	0.21 *	0.60 ***	0.45 **	0.50 **	0.46 **
	Basic treatment group				**Control group**			
	AoI-A	AoI-E	AoI-T	AoI-S	AoI-A	AoI-E	AoI-T	AoI-S
AoI Artifact	1.00 ***	0.57 ***	0.60 ***	0.45 **	1.00 ***	0.40 **	0.60 ***	0.07 n.s.
Survey - Use 1	0.64 ***	0.44 **	0.36 *	0.25 n.s.	0.00 n.s.	0.02 n.s.	0.07 n.s.	-0.08 n.s.
Survey - Use 2	0.59 ***	0.38 *	0.36 *	0.21 n.s.	0.03 n.s.	0.13 n.s.	0.12 n.s.	0.01 n.s.

AoI-A = AoI artifact | AoI-E = AoI email client | AoI-T = AoI ticket system |
AoI-S = AoI service catalog
*** p < 0.001 | ** p < 0.01 | * p < 0.05 | n.s. p > 0.05

The test reveals a correlation between the treatments' objective usage (AoI-Treatment) and the subjective usage measurements (Use1 and Use2) for the complete dataset, the "advanced treatment" group, and the "basic treatment" group. The correlation test for the control group reveals no correlation between the perceived and objective usage measurements. In addition, we tested for a correlation between the subjective usage and the objective usage of the experimental applications. In line with the first correlation test, we found a positive correlation for the complete data set and both treatment groups (except for the "basic treatment" group's service catalog usage).

There is no correlation for the control group. As last step, we tested for a correlation between the objective artifact usage and objective usage of the experimental applications. Here, we found a positive correlation for the complete dataset as well as all three experimental groups (except for the control group's service catalog usage).

6 Discussion

The analysis of the complete data set reveals a positive correlation between the subjective usage perceptions and the objective artifact usage. Moreover, we identified a correlation between the subjective artifact usage and the objective experimental applications usage. This could indicate a type II error, as the participants might not only have assessed their artifact usage, but the overall usage of the experiment's applications. Therefore, we tested for a correlation between the objective artifact's and experimental application's usage. Again, we found a positive correlation. The experiment's setup and task could explain these correlations. The participants used the experimental applications in order to fulfill their experimental tasks and the assistance system supported their process execution. Therefore, there is a correlation between all four objective usage measurements. This fact could also explain the correlation between the subjective artifact's usage and objective usage of the experiments applications. Based on the introspection theory [17], we assume the participants are able to assess their artifact usage, as their cognitive artifact usage is stored in their short-term memory. The objective usage measurements' reliability is based on the immediacy assumption and the eye-mind assumption [12]. Therefore, we can conclude both measurements as reliable and confirm our hypothesis for the entire dataset.

However, the distinct analysis for the three groups reveals contradictious results. Both treatment groups show similar results compared to the complete dataset. However, the control group reveals a different result. There is no correlation between the subjective and objective artifact usage. The control group participants had no information about how to execute the ticketing process provided in their artifact. Thus, we expected that they do not use it (cognitively) in the experiment as their artifact contained only a blank space. Thus, we assume they have the information on "no artifact usage" in their short-term memory. Accordingly, the control group's participants should not report any usage. Nevertheless, the control group reported a low subjective artifact usage (mean value of 2.76 and 2.54).

There are some possible explanations, for example common method biases such as the social desirability bias [14, 35]. The control group participants may had the tendency or need to report some artifact usage in order to "*present themselves in a favorable light*" [14: p. 881], despite the fact they actually did not use it. The participants of the control group could have over-reported their usage because they viewed this behavior as appropriate [35]. Another explanation could be the control group participants were confused and reported their subjective usage of the experimental applications. However, the analysis showed there is no such correlation and therefore, we assume there was no confusion among the participants. Summed up, we can confirm the hypothesis for both treatment groups and reject the hypothesis for the control group.

7 Summary and Future Work

In this paper, we present our research addressing the ongoing debate on subjective vs. objective usage measurements. More specifically, we address experimental settings with cognitive artifact usage, e.g. in the context of DSR projects. In our experiment, we measure the objective usage with eye-tracking technology and the participants' eye fixations. Our analysis supports the assumption that subjective usage measurements are a reliable approximation for objective usage measurements in case of laboratory experiments with an artifact present. When testing a group with an artifact against a group without an artifact the subjective measurements are no reliable approximation for objective measurements according to our analysis. We conclude that researchers could use subjective usage measurements when comparing different artifact designs. In contrast to this, researchers should rely on objective measurements when testing the effect of an artifact compared to a control group with no artifact.

Our research contributes by providing support for the constraint reliability of subjective usage measurements in experimental settings discussed in the body of knowledge. More specifically, our research results indicate the reliability of participants' usage perceptions of an artifact in an experimental setting in case if the artifact is present and can be (cognitively) used by the participants. Moreover, from a methodological point of view, we exemplary show how to measure objective cognitive artifact usage in experimental settings, based on the immediacy assumption and the eye-mind assumption [12]. Eye-tracking technology enables researchers to utilize the participants' eye fixation as a reliable indicator for their cognitive processing and thus, represents an objective usage measurement. Researchers can apply the presented method to assess the participants' (cognitive) artifact usage in their experiments. Especially experiments in the DSR context can benefit from using eye-tracking technology to enrich the usage measurement [15] as the artifact usage is of high importance in DSR experiments [10]. Researchers following the DSR approach can apply the proposed objective usage measurement in their experiments aiming to gain deeper insights into the participants' usage of the artifact and the artifacts' effects.

Although our research follows established methods, there are potential limitations. We conducted the analysis with only one experimental dataset resulting in limited generalizability. There is an ongoing debate on subjective vs objective usage measurements and our findings add to the body of knowledge. Moreover, our study presents a reliable method on how to assess objective measurements for cognitive artifact usage in experimental settings. The analysis presented in this paper uses the aggregate experimental data as our research is still in progress. As next step, we conduct a detailed analysis of the participants' individual usage behavior during the experiment addressing the following activities. **First**, the screen recordings from the eye-tracking software enables us to track the participants' activities and fixations during the experiment, code the participant's usage behavior accordingly, and derive patterns of varying usage behaviors. **Second**, for the presented analysis, we used the total numbers of fixations as the objective measurement. As the used eye-tracking technology measures every 33ms the participant's eye position on the screen, the gathered fixations might also include fixations that occurred, because the participant is scanning the screen (e.g. investigating the applications in the beginning). Based on reading research in psychology, we plan to code fixation clusters with a duration of at least 200 – 300ms, as this time is required to read and comprehend a sentence [25]. This detailed analysis will enable us, for example, to remove the "non-usage" fixations in the experiment's beginning, when the participants scanned the existing applications, in order to get a clearer picture of their actual usage behavior during the experiment. Both research activities will extent the "richness" of our usage measurement following the suggestions by Burton-Jones and Straub [15]. **Third**, we will incorporate the artifact's usage measurements to explain the participants' actual performance in the experiment and test for moderating variables such as gender, ethnicity, and users' knowledge about the experiments ticketing process. Especially the varying level of participants' process knowledge, which increased during the experiment based on the provided assistance, might have an effect on the participants' usage behavior. The detailed analysis of the eye-tracker recordings will enable us to understand the participants changing usage behavior during the experiment and gain further insights on potential learning effects based on the provided assistance.

In addition to the outlined activities addressing our future research, there are further opportunities. We suggest to conduct more studies on the proposed concept of cognitive artifact usage. There is more work required on the theoretical grounding and conceptualization of cognitive artifact usage. In addition, researchers could apply the eye-tracking technology in their experiments to gather objective usage data and provide more findings on the ongoing subjective vs objective usage debate.

References

1. Lucas, H.C., Spitler, V.K.: Technology Use and Performance: A Field Study of Broker Workstations. Decision Science 30, 291–331 (1999)
2. Devaraj, S., Kohli, R.: Performance Impacts of Information Technology. Is Actual Usage the Missing Link? Management Science 49, 273–289 (2003)
3. Eckhardt, A., Maier, C., Hsieh, J.P.-A., Chuk, T., Chan, A.B., Hsiao, J.H., Buettner, R.: Objective measures of IS usage behavior under conditions of experience and pressure using eye fixation data. In: ICIS 2012 Proceedings (2012)
4. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 13, 319–340 (1989)
5. Reuver, M. de, Bouwman, H.: Dealing with self-report bias in mobile Internet acceptance and usage studies. Information & Management 52, 287–294 (2015)
6. Barnett, T., Kellermanns, F.W., Pearson, A.W., Pearson, R.A.: Measuring Information System Usage: Replication and Extensions. Journal of Computer Information Systems 47, 76–85 (2006)

7. Straub, D., Limayem, M., Karahanna-Evaristo, E.: Measuring System Usage. Implications for IS Theory Testing. Management Science 41, 1328–1342 (1995)
8. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28, 75–105 (2004)
9. Bhattacherjee, A.: Social Science Research: Principles, Methods, and Practices. Textbooks Collection. Book 3, http://scholarcommons.usf.edu/oa_textbooks/
10. Mettler, T., Eurich, M., Winter, R.: On the Use of Experiments in Design Science Research: A Propositionof an Evaluation Framework. Communications of the Association for Information Systems 34 (2014)
11. Gregor, S., Benbasat, I.: Explanations from Intelligent Systems: Theoretical Foundations and Implications for Practice. MIS Quarterly 23, 497–530 (1999)
12. Just, M.A., Carpenter, P.A.: A theory of reading: from eye fixations to comprehension. Psychological Review 87, 329–354 (1980)
13. Wu, J., Du, H.: Toward a better understanding of behavioral intention and system usage constructs. Eur J Inf Syst 21, 680–698 (2012)
14. Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P.: Common method biases in behavioral research: a critical review of the literature and recommended remedies. The Journal of applied psychology 88, 879–903 (2003)
15. Burton-Jones, A., Straub, D.W.: Reconceptualizing System Usage. An Approach and Empirical Test. Information Systems Research 17, 228–246 (2006)
16. Nisbett, R.E., Wilson, T.D.: Telling more than we can know. Verbal reports on mental processes. Psychological Review 84, 231–259 (1977)
17. Ericsson, K.A., Simon, H.A.: Verbal reports as data. Psychological Review 87, 215–251 (1980)
18. Dulany, D.E., O'Connell, D.C.: Does partial reinforcement dissociate verbal rules and the behavior they might be presumed to control? Journal of Verbal Learning and Verbal Behavior 2, 361–372 (1963)
19. Bryce, D., Bratzke, D.: Introspective reports of reaction times in dual-tasks reflect experienced difficulty rather than timing of cognitive processes. Consciousness and Cognition 27, 254–267 (2014)
20. Bratzke, D., Bryce, D., Seifried-Dübon, T.: Distorted subjective reports of stimulus onsets under dual-task conditions. Delayed conscious perception or estimation bias? Consciousness and Cognition 30, 36–47 (2014)
21. van Gog, T., Scheiter, K.: Eye tracking as a tool to study and enhance multimedia learning. Learning and Instruction 20, 95–99 (2010)
22. Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S.W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., Tsai, C.-C.: A review of using eye-tracking technology in exploring learning from 2000 to 2012. Educational Research Review 10, 90–115 (2013)
23. Jacob, R.J., Karn, K.S.: Eye Tracking in Human-Computer Interaction and Usability Research. In: Hyönä, J., Radach, R., Deubel, H. (eds.) The mind's eye. Cognitive and applied aspects of eye movement research, pp. 573–605. North-Holland, Amsterdam, Boston (2003)
24. Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychological Bulletin 124, 372–422 (1998)

25. Rayner, K.: Eye movements and attention in reading, scene perception, and visual search. Quarterly journal of experimental psychology 62, 1457–1506 (2009)
26. Traxler, M.J., Long, D.L., Tooley, K.M., Johns, C.L., Zirnstein, M., Jonathan, E.: Individual Differences in Eye-Movements During Reading: Working Memory and Speed-of-Processing Effects. Journal of eye movement research 5 (2012)
27. Marti, S., Bayet, L., Dehaene, S.: Subjective report of eye fixations during serial search. Consciousness and Cognition 33, 1–15 (2015)
28. Johansson, P., Hall, L., Sikstrom, S., Olsson, A.: Failure to detect mismatches between intention and outcome in a simple decision task. Science (New York, N.Y.) 310, 116–119 (2005)
29. Maedche, A., Morana, S., Schacht, S., Werth, D., Krumeich, J.: Advanced User Assistance Systems. Bus Inf Syst Eng (2016)
30. Morana, S., Schacht, S., Scherp, A., Maedche, A.: Designing a Process Guidance System to Support User's Business Process Compliance. In: ICIS 2014 Proceedings (2014)
31. Tobii: About Tobii Pro, http://www.tobiipro.com/about/
32. Royston, J.P.: An Extension of Shapiro and Wilk's W Test for Normality to Large Samples. Applied Statistics 31, 115 (1982)
33. Hollander, M., Chicken, E., Wolfe, D.A.: Nonparametric statistical methods. John Wiley & Sons, Inc, Hoboken, New Jersey (2014)
34. Best, D.J., Roberts, D.E.: Algorithm AS 89. The Upper Tail Probabilities of Spearman's Rho. Applied Statistics 24, 377 (1975)
35. Donaldson, S.I., Grant-Vallone, E.J.: Understanding Self-Report Bias in Organizational Behavior Research. Journal of Business and Psychology 17, 245–260 (2002)

How Does the Crowdsourcing Experience Impact Participants' Engagement? An Empirical Illustration

Julia Troll[1], Ivo Blohm[1]

[1] University of St. Gallen (HSG), Institute of Information Management, St. Gallen, Switzerland
julia.troll,ivo.blohm@unisg.ch

A largely neglected aspect in crowdsourcing research is the "Crowdsourcing Experience" itself, which every crowdsourcee is necessarily exposed to throughout the IT-mediated interaction process, potentially stimulating engagement towards the crowdsourcer. Hence, the crowdsourcees' engagement process is conceptualized and illustrated with empirical findings from a pilot case. It exemplifies that crowdsourcing has the potential to generate high levels of attitudinal and behavioral engagement, depending on prior experiences and perceived cognitions and emotions. Related stimuli characteristics are identified, which serve as a first indication of the foundations of the engagement process. This study offers IS-researchers first insights on the so far under-researched topic of IT-enabled engagement processes between individuals and entities.

Keywords: Crowdsourcing, Crowdsourcing Experience, Customer Engagement, Engagement Process

1 Introduction

Crowdsourcing is an emerging global trend, which 85 percent of the top hundred global brands try to take advantage of [1]. It broadly defines a participative, IT-mediated activity in which a given entity proposes a task to a crowd to create mutual benefit [2, 3]. While there are several functions of crowdsourcing, such as design and innovation, or software development and testing [4], it seems as if crowdsourcers' primary attention is currently paid to managing contributions rather than the crowd, its needs and desires. This is also reflected by research in the field of crowdsourcing, which is dominated by studies assessing crowdsourcing mostly from a crowdsourcer's perspective. However, looking at successful crowdsourcing initiatives, as *My Starbucks Idea* or the *SBB Mobile Preview Community*, in terms of its huge crowds and intense participation, it can be assumed that value is not only created by absorbing knowledge and ideas.

The meaning of value and the process of value creation are rapidly shifting from a product- and firm-centric view to an experience-based view, putting the subject in the center [5]. This can be transferred to co-creation activities itself, in which experiences

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Troll, J.; Blohm, I. (2017): How Does the Crowdsourcing Experience Impact Participants' Engagement? An Empirical Illustration, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1036-1050

are created, too. Hence, this paper argues that a largely neglected aspect in crowdsourcing research is the here called "*Crowdsourcing Experience*" itself, which every crowdsourcee is necessarily exposed to throughout the IT-mediated interaction process. This disregard may come with a price. Initiators not only risk to lose valuable contributors during or after the interaction due to perceived negative experiences, but also their reputation. A famous example is given by *Pril's crowdsourcing flop*, in which an undesirable experience by Henkel caused a public PR-disaster [6]. Additionally, initiators miss a promising opportunity to generate crowdsourcees a unique experience, thereby stimulating overall engagement towards the crowdsourcer. This can create additional value, e.g., in form of positive word of mouth and enhanced brand value, increasing in relevance if the crowd consists of (potential) customers and end-users.

First authors recognized the need for an experienced-based perspective on crowdsourcing and called for research [4, 7, 8]. However, no existing study takes a process perspective to systematically assess the end-to-end crowdsourcee's experience. Yet, this is necessary to understand how and why crowdsourcees engage, from a cognitive, emotional and behavioral perspective. To fill this gap, the engagement process is conceptualized, illustrated, and refined with empirical observations from a case, to approach the following question: How does the *Crowdsourcing Experience* impact engagement throughout the IT-enabled interaction process?

First an overview of the research field of crowdsourcing and customer engagement is provided and relevant concepts derived. Then, an empirical illustration is provided and a refined concept discussed. Lastly, relevant research contributions are presented.

2 Conceptual and Theoretical Background

2.1 Crowdsourcing

The fundamental idea of crowdsourcing is that a crowdsourcer (e.g., a company) proposes to an undefined group of contributors (e.g., individuals), henceforth called crowdsourcees, the voluntary undertaking of a task presented in an open call [2]. The ensuing interaction process unfolds over IT-based crowdsourcing platforms [2, 3]. Crowdsourcers can set up their own crowdsourcing platform and processes (e.g., My Starbucks Idea), or they can refer to intermediaries, such as Innocentive or Testbirds that provide a technical infrastructure and access to a crowd. Some offer additional services such as task specification, crowd acquisition, and evaluation of results [9]. Crowdsourcer and crowdsourcees engage in the participative, IT-mediated interaction process to create mutual benefit [3]. For crowdsourcers, this benefit may involve solving problems that cannot be satisfactorily solved in-house, but also enhanced brand visibility [10]. For crowdsourcees, the benefit may be of economic nature (e.g., remuneration) or other needs are satisfied, like social recognition or skill development. Thus, value can be produced by outcomes (i.e., instrumental value) and preceding processes (i.e., experiential value). To better understand the mutual benefits of crowdsourcing, some authors have emphasized the need for researching crowdsourcing from an experience-based perspective [4, 7, 8]. First articles reveal insights on: initial crowdsourcing user engagement, defined as the quality of effort [11]; drivers of sustained participation in micro-task oriented crowdsourcing [12]; an behavioral engagement index for crowdsourcing [13]; crowdsourcee's attitude towards the platform and design choices [14]; and the impact of crowdsourcing on affective commitment in collaborative crowdsourcing projects [15]. It seems that each of those studies either focus on a specific crowdsourcing phase in the interaction process or solely on the experience outcome, from a behavioral or attitudinal perspective. None of those studies take a holistic process perspective to systematically assess the end-to-end crowdsourcees' experience, including pre- and post-participation experiences. Yet, this is necessary to understand how and why crowdsourcees engage for value co-creation. This paper takes a closer look at the concept and process of customer engagement from the relationship marketing literature and applies it to crowdsourcing.

2.2 The Concept and Process of Customer Engagement

Customer engagement (CE) is defined as a psychological state that occurs by virtue of interactive, co-creative customer experiences with a focal agent [16]. Customer experience is the internal and subjective perception of customers' direct and indirect interactions with a firm. The resulting engagement state develops through a dynamic, iterative process that co-creates value between the engagement subject (e.g., customer) and object (e.g., company) [17]. First authors conceptualized the general engagement process of customers [18, 19]. A simplified illustration is given in Figure 1.

According to existing conceptualizations, the psychological state encompasses various combinations of cognitive, emotional, and behavioral dimensions, dependent on perceived stimuli and prior experiences. The cognitive dimension can be interpreted as a more passive state of immersion and absorption or a more active state of cognitive processing to expedite comprehension [20]. The emotional dimension relates to the customer's feelings activated by an experience. Additionally, a behavioral response related to a specific stimulus may be expressed. Addressed dimensions regarding each perceived stimulus are evaluated by the subject and an intermediate state is generated, happening unconsciously. The literature considers satisfaction, delight, involvement and trust as intermediate states that foster the development of engagement, which is defined as a specific type of commitment towards the engagement object [18, 19].

Satisfaction is generally seen as a preliminary state. Alone, it may not result in a desired behavior (i.e., repeat consumption or referral) as expectations are only confirmed according to expectation-disconfirmation theory [21]. If one repeats a satisfying interaction due to perceived attribute-based utility, missing alternatives or switching costs, a so called calculative commitment may develop between the engagement subject and object [19]. Commitment is associated with a specific attitudinal position [19], while calculative relates to rational reasoning. However, those rational bonds may be dissolved easily and are of limited value for a company [18].

Hence, additionally an emotional bond is desired, also called affective commitment. It illustrates a customer's psychological closeness to a focal agent and is positively related to referral and word of mouth (WOM) [18, 19]. It is expressed as a holistic or aggregate judgment, independently from its functional attributes. A feeling of involvement or trust, due to increased familiarity and precise expectations towards the engagement object, is known as a driver [19]. While involvement is described as a feeling of personal relevance and importance, trust is a customer's assumption that a focal agent is able to respond to his needs and has his best interest at heart [22]. Nevertheless, a delightful incident may lead to affective commitment right away, even if the engagement subject is less familiar with the engagement object and relies on a more attribute-based evaluation [18, 19]. Customer delight is defined as a combination of pleasure, joy and elation as well as unexpected levels of arousal or surprise [23].

When both forms of commitment develop throughout the interaction process, customer and company are in an enduring relational exchange with strong emotional bonds [18]. This desired psychological engagement state is related to direct (i.e., repeat consumption) as well as indirect behavioral responses towards the engagement object (e.g., WOM, referral behavior), reflecting the customer engagement value [24].

It can be concluded that familiarity (i.e, prior experiences) with an engagement object is an input factor in the process of engagement, while the experience evaluation constitute the psychological process, leading to a state of calculative and/or affective commitment and behavioral responses as process outcomes. Presuming a feeling of satisfaction, it is supposed that a sense of delight, involvement, and trust operate as drivers of engagement in a customer-company interaction. Active participation in the creation of an offering is widely assumed as a central antecedent [16, 17, 24, 25].

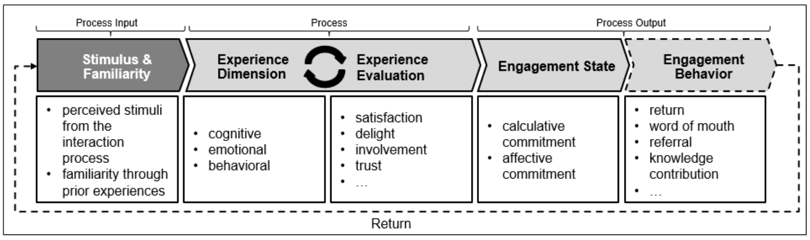


Figure 1. Conceptualized Engagement Process (own illustration)

3 Towards an Engagement Theory of the Crowdsourcing Experience

Independent of the crowdsourcer's original intention, performing a crowdsourcing initiative creates an experience that may foster engagement among crowdsourcees towards the crowdsourcer. The *Crowdsourcing Experience* in this paper is defined as a crowdsourcee's internal and subjective perception of the end-to-end, IT-mediated interaction process, resulting in a psychological state. It is an online experience, driven by several stimuli over one or more virtual channels. Perceived stimuli can be found in the pre-participation (e.g., invitation), participation (e.g., task), and post-participation (e.g., payment) phase. Due to its participative character, the underlying assumption is that crowdsourcing generally has the potential to generate high levels of engagement. Depending on the specific set up of the initiative, crowdsourcees (i.e., the engagement subject) may engage with the crowdsourcer directly or via an intermediary and with other crowdsourcees (i.e., the engagement objects). They can have varying degrees of familiarity concerning the objects (e.g., prior crowdsourcing- or customer experiences), influencing their expectations and experience evaluation. Henceforth, *Crowdsourcee Engagement* is conceptualized as a psychological process that models the underlying mechanisms by which a crowdsourcee develops calculative and affective commitment based on perceived stimuli and prior experiences, resulting in behavioral value-contributions for the crowdsourcer. The unfolding IT-mediated interaction process comprises a set of diverse stimuli, potentially addressing both, the cognitive and emotional experience dimension. The CE literature considers satisfaction, delight, involvement and trust as intermediate states. To explain potential drivers in the context of crowdsourcing, different perspectives can be taken, as crowdsourcees may not only be seen as (potential) customers and influencers, but also take the role of a platform user, worker, and a group or community member (i.e., the crowd).

From an IS-perspective, a system's characteristics, quality and performance may generate user involvement, delight, and trust. For example, characteristics as novelty, variety, aesthetics (affective or sensory appeal), and fun are related to perceived delight [26, 27]. In crowdsourcing, this may refer to an attractive and fun-providing crowdsourcing platform or an appealing virtual object, which is in the center of the task (e.g., a website). According to organizational behavior (OB) research, specific task characteristics, one's identity with it, and rewards may lead to job or task involvement, trust or delight [28]. For example, a good task-person fit and a crowdsourcee's enthusiasm about a task may be related to involvement and delight. Lastly, according to community research, the identification with the crowd may stimulate a sense of involvement throughout the process [29]. Next to these, another driver of engagement is expected to operate in the case of crowdsourcing: empowerment. Ulrich [30] argues that customer empowerment leads to stronger commitment, if additional information about the company can be gained and response is volitional, irreversible, and public. Empowerment positively effects demand and WOM, due to a sense of psychological ownership [31]. In OB-research, it relates to a sense of control, impact, meaning, and self-efficacy [32], which may be stimulated e.g., with a specific task.

Subsequently, out of the intermediate states an overall engagement state arises. If satisfaction is achieved and the crowdsourcee perceives clear utility through participation, a form of calculative commitment towards the crowdsourcer may be gained. If additionally to satisfaction, a sense of delight, involvement, trust, and/or empowerment arises throughout the interaction process, affective commitment may be developed. Resulting direct and indirect behavioral value contributions towards the crowdsourcer may refer to: a) repeat participation; b) virtual or direct WOM; c) referral behavior; d) further voluntary knowledge or feedback contributions, exceeding the scope of the original task; as well as e) consumption activities (buying/ using something from the crowdsourcer). Calculative commitment is related to repeat participation (a) and affective commitment additionally to indirect contributions (b-e).

By assessing the engagement process in the context of an exemplary crowdsourcing case, those relationships will be illustrated and successful patterns of mechanisms and related stimuli characteristics extracted to refine and extend derived knowledge.

4 An Empirical Illustration

Each crowdsourcing initiative can offer crowdsourcees a unique IT-mediated interaction process, consisting of many consecutive and interrelated experience-driving stimuli. This section illustrates how the concept and process of engagement can be useful for interpreting the findings of a qualitative study that investigated the perceived *Crowdsourcing Experience* of participants in a crowdsourcing project, initiated by a leading insurance company from Switzerland. This approach is accepted by recognized outlets and a successful example is provided by Leonardi [33].

4.1 Case Description

In 2015, InsureCorp (name changed) decided to renew its digital communication channels with a “mobile first” strategy. To apply a user-centered approach for developing its new mobile web application, the company decided to use crowdsourcing with potential end-users. Crowdsourcees were offered to test and feedback the web app’s interface and report on functional bugs, usability and provide ideas. They had to go through realistic test scenarios to explore the web app. In return, they were offered a fixed monetary reward. InsureCorp chose to cooperate with a crowdsourcing intermediary, responsible for acquiring the crowd, providing the platform, evaluating contributions, and handling the payment process. They conducted three self-contained crowdsourcing projects (August 2015; January and June 2016), each with a duration of five days, to individually advance parts of the web app with around twenty crowdsourcees per iteration. Each project included the acquisition of a suitable crowd, a definite task, and a closing phase. The last project was assessed in this study.

The case of InsureCorp was chosen because it illustrates a common case in this field and incorporates all characteristics of crowdsourcing, as a concrete task is proposed via an open call through a platform for a specified reward. The goal was to target a diverse crowd, representing potential end-users. As the company developed a certain maturity over iterations, it is expected that in the last one exceptional problems, unusually influencing the *Crowdsourcing Experience*, could be reduced. The crowd was relatively homogenous regarding cultural background, familiarity with the activity, and financial situation, which enabled a comparison of experiences and engagement processes. Lastly, the use of intermediaries is becoming increasingly common [9]. Hence, it could be explored in how far the engagement of crowdsourcees developed differently towards the crowdsourcer, as the central point of interest, and the intermediary.

4.2 Data Collection and Analysis

First, to understand the intended *Crowdsourcing Experience*, three semi-structured interviews and a focus group discussion with the crowdsourcer and intermediary were conducted. Also, to study crowdsourcees’ behavior, data concerning the time spent on the platform and with the web app was tracked. Contributions were analyzed in terms of its length (word count) and level of detail (i.e., under-/ over-fulfillment of task). Demographic information and amount of previous activities were collected from the platform. Finally, seven in-depth, semi-structured interviews (60-90 minutes) with crowdsourcees were conducted to decipher the crowdsourcee's experience along the process. A slightly adapted version of the novel approach from consumer behavior, called "Sequential Incident Laddering Technique" (SILT), was used [34]. Respondents were first asked to recall all stimuli ("critical incidents") from the crowdsourcing interaction process. Subsequently, the interviewer asked simple "what", "why", "how" questions to establish the link between a stimulus and crowdsourcee's (a) cognitive and emotional perceptions; (b) experience evaluation (intermediate state); (c) and behavioral responses ("laddering technique"). In a last interview step, crowdsourcee's final commitment and (planned) engagement behavior towards the crowdsourcer and intermediary was captured. As commitment is also described as an attitudinal judgment, interviewees were asked to describe their attitude to receive insights regarding their emotional and rational disposition. To avoid a recall bias [35] crowdsourcees in this study were interviewed two to seven days after participation. For reasons of better comparability, seven crowdsourcees with some crowdsourcing familiarity were selected, to avoid interviewing overly excited or bored individuals. The interviews were transcribed and assessed, together with the other data sources, by applying qualitative content analysis [36, 37]. A category system based on the theoretical framework of the engagement process was developed and collected data was coded along stimuli: perceived experience dimensions; related engagement object; experience evaluations; resulting attitude; and (planned) behavior. To allow for the identification of new categories and related stimuli characteristics, the system was iteratively adapted. Two researchers independently coded the data by allocating direct and indirect statements to the categories (interpretive approach) and subsequently discussed findings. Insights were used to illustrate how engagement developed for those crowdsourcees throughout the process. The purpose was not to test the framework, but rather to illustrate its use for understanding the potential engagement value of a crowdsourcing initiative.

4.3 The Crowdsourcing Interaction Process of InsureCorp

First, potential experience-driving stimuli along the interaction process were visualized based on the results of the interviews with responsible project managers. The process was then collaboratively discussed and refined in a focus group interview. The result is illustrated in Figure 2.

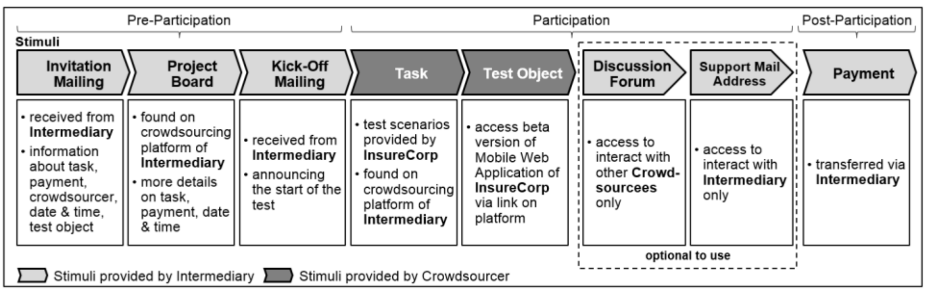


Figure 2. Stimuli along the Crowdsourcing Interaction Process of InsureCorp

It includes two communication channels: email and the crowdsourcing platform of the intermediary. Three potential engagement objects could be identified: crowdsourcer, intermediary, and other crowdsourcees. Five stimuli are solely designed, managed, and communicated by the intermediary to the crowd, while two stimuli (task, test object) are designed and managed by the crowdsourcer. One stimulus (discussion forum) is provided by the intermediary but triggers the interaction among crowdsourcees only.

4.4 Assessment of the Crowdsourcing Experience

By looking at the described attitudes and (planned) behavior, it is observed that different engagement states among crowdsourcees developed, although the overall *Crowdsourcing Experience* was evaluated to be satisfying for all crowdsourcees. A more in-depth analysis of the underlying processes was necessary to identify mechanisms that caused psychological and behavioral engagement outcomes.

The case data shows that the engagement development process throughout the interaction process took several forms among crowdsourcees, depending on prior familiarity with the engagement objects and stimuli perceptions. All previously identified stimuli were generally perceived and mostly experienced by interviewed crowdsourcees, shaping their *Crowdsourcing Experience*, except from the discussion board. That excludes “other crowdsourcees” as a potential engagement object. As expected, crowdsourcees related the stimuli task and test object directly to the crowdsourcer and the rest to the intermediary. Stimuli, related to the intermediary, lead to 80 percent cognitive statements and 54 percent were evaluated to be purely satisfying, while 25 percent were additionally evaluated as trust-enhancing. Stimuli related to the crowdsourcer lead to around 60 percent cognitive statements and only 17 percent were evaluated to be purely satisfying, while even 73 percent were additionally evaluated as delighting, involving, or empowering (see appendix for more details). An integrated framework, incorporating theoretical knowledge from the engagement process (Figure 1) with findings from the assessment of InsureCorp’s crowdsourcing interaction process (Figure 2), is illustrated in Figure 3.

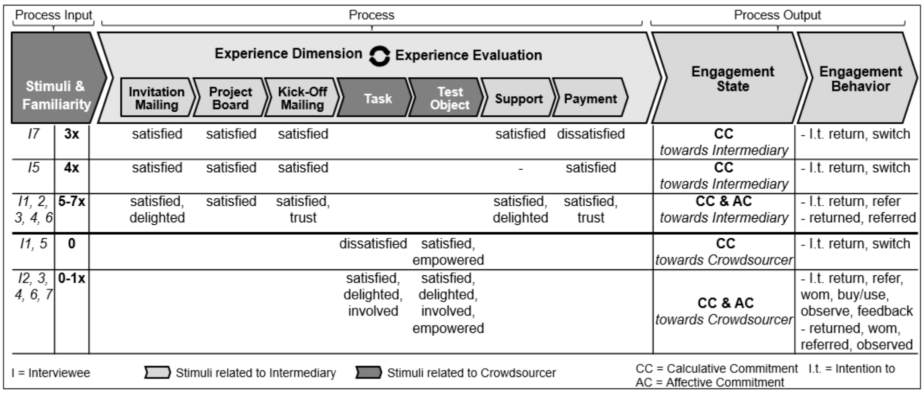


Figure 3. Integrated Framework of the Crowdsourcee’s Engagement Process

Different process patterns were identified that led to affective commitment towards the crowdsourcer and intermediary. First, those five interviewed crowdsourcees (I2, 3, 4, 6, 7), who evaluated the stimuli task and test object as delightful, involving, and/ or empowering, developed a more emotionally based commitment towards the crowdsourcer, leading to more diverse behavioral responses than the other two. Delight arose due to a feeling of surprise, pleasure, fun and enjoyment related to the task and test object (e.g., I2: “it was fun to explore the whole web app and record my feedback in a video”; I7: “those scenarios were new to me, I felt like a real customer”). A feeling of challenge, inspiration, stimulation, and need for solving the task was mentioned when crowdsourcees described themselves as being involved (e.g., I3: “I couldn’t find it but I really wanted to solve that task, so I tested the whole application”; I6: “the app design was very inspiring, it was easy to get caught up by the task”). Those, who felt as being a part of the product-development process and enjoyed having impact on the test object, described the stimulus as empowering (e.g., I2: “it feels good to give feedback for a product that is still in development”). Crowdsourcees, who described those emotional perceptions and experience evaluations, stated that their attitude towards the crowdsourcer changed somewhat, as they perceived InsureCorp as more innovative, modern, open-minded, collaborative, customer-centric, and/or supportive after participation. They also mentioned an improved brand image and a strengthened relationship to the crowdsourcer (e.g., I4: “now, InsureCorp feels more like a partner for me”; I2: “I did not expect that from InsureCorp, seems like a cool company”). This indicates a sign of a stronger form of affective commitment towards the crowdsourcer. While all interviewed crowdsourcees stated to be generally willing to return for repeat participation based on perceived utility, those that mentioned to be delighted, empowered and/ or involved concerning task and test object, were additionally intended to refer the crowdsourcer, conduct of WOM, and buy or use a service of the crowdsourcer due to their positive impression after participation. Some were also interested in observing the development of the test object and providing voluntarily, additional feedback and ideas to the crowdsourcer after the project’s official end. Data showed that involved crowdsourcees spent more time on the platform and with the test object in comparison to others (1.5 to 2 times as long). Contribution-analysis revealed that they did more than was expected in the task (over-fulfillment) and gave more detailed feedback in terms of word count (1.25 to 1.6 time as much). In comparison, those that perceived only satisfaction or even dissatisfaction regarding the task and test object mentioned no intentions for referral, WOM, consumption or observation towards the crowdsourcer and contributed less in terms of feedback.

Second, those five crowdsourcees (I1, 2, 3, 4, 6), who evaluated stimuli related to the intermediary mainly as satisfying but expressed that they developed some trust into the intermediary throughout the process, developed some affective commitment, resulting in more diverse behavioral value contributions for the intermediary. Satisfaction with stimuli as the invitation mailing, project board, and support services, was mainly described through cognitive expressions, relating to the characteristics of the information provided, the platform, or the response time. A feeling of trust towards the intermediary was mentioned in relation to the kick off mailing (e.g., I4: “I felt relieved, when the reminder arrived. I know, I can rely on their processes”), and the compensation (e.g., I2: “I don’t know what others pay, but I assume they are fair”; I1: “the process could be easier, but I’m sure they`ll find a better solution soon”). Negative cognitions of crowdsourcees, who mentioned trust into the intermediary, resulted not in negative emotional perceptions and evaluations. In comparison, other crowdsourcees expressed annoyance in response, resulting in dissatisfaction. Although crowdsourcees described their attitude towards the intermediary mostly rational and used terms as responsive, fair, reliable, effective and well-organized, those that sensed trust throughout the process, used more emotional expressions for the intermediary (e.g., I4: “it was fun to work with them”; I6: “they try their best to make our job easier”) and were willing to refer (or even already referred) the intermediary to friends or colleagues. In comparison, those that mentioned only satisfaction or even some dissatisfaction without showing signs of trust, were only intended to return due to rational reasons of perceived utility (e.g., compensation and skill development), but mentioned to be willing to switch, if another crowdsourcing opportunity arises (e.g., I5: “the intermediary is for me more a means to an end”; I7: “I don’t have any emotional relationship with it”). Thus, only a calculative commitment can be assumed.

The difference between the development of affective commitment towards the crowdsourcer and intermediary may be explained due to two reasons. On the one hand, prior familiarity may play a role. All interviewed crowdsourcees were already familiar with the intermediary (three to seven prior projects) and those that developed trust participated in five to seven other crowdsourcing projects before. Hence, they had quite precise knowledge and expectations regarding the general interaction points, designed and managed by the intermediary. Instead of being easily surprised (i.e., delighted), they rather valued repetitions and dependability, which enhanced their trust. In comparison, familiarity with the crowdsourcer was much lower. Only two crowdsourcees participated in one of the previous iterations. Thus, most were more sensitive for positive surprises. On the other hand, the type of stimuli, related to the intermediary were much less involving or empowering and more of an administrative character, than the ones related to the crowdsourcer. The task and test object allow for intense interaction with the crowdsourcer than a rather transactional stimulus, as an informative mail or payment process. From cognitive and emotional stimuli perceptions and evaluations, relevant characteristics could be identified (see Table 1). Derived characteristics illustrate the foundations of the crowdsourcee’s engagement process.

Table 1. Perceived Stimuli Characteristics

Intermediate State	Related Stimuli Characteristics (as perceived by interviewed Crowdsourcees)
Satisfaction	(a) complete, concrete, understandable information and instructions, (b) clear in/ out-of-scope of task, (c) easy to use crowdsourcing-platform, (d) easy access to test-object (e.g., easy registration, technol. prerequisites), (e) quick response time for support, (f) monetary compensation, (g) quick compensation transaction
Delight	(a) personal style of contact (e.g., personal address, real contact person as sender) (b) personal/direct communication channel for invitation/support (e.g., email/ phone), (c) new/innovative type of task (e.g., video feedback), (d) explorative task (e.g., usability testing), (e) fun-providing test scenarios, (f) new/ innovative design of test object

Involvement	(a) challenging task, (b) stimulating and inspiring design of test-object/ information provided, (c) realistic test scenarios (e.g., put them in the position of a real customer)
Trust	(a) process transparency (e.g., comprehensive information through reminder mails, process details on platform, regular updates), (b) fair compensation (in terms of time and effort), (c) process improvement-attempts/ actions (e.g., news announcing changes)
Empowerment	(a) having impact on whole test object (e.g., explorative task, broad scope of task), (b) changes/ developments in test object at project-end (e.g., feedback report)

5 Discussion

This paper began with the suggestion to take a holistic process perspective for systematically assessing the end-to-end *Crowdsourcing Experience* to understand how and why crowdsourcees actually engage for value co-creation. Therefore, in analogy to the CE-process, a theoretical engagement process for the case of crowdsourcing was derived and its use illustrated with a case.

The underlying assumption was that crowdsourcing generally has the potential to generate high levels of engagement due to its participative character [16, 17, 24, 25]. The attitudinal and behavioral responses by participants in the case illustrated that emotional as well as rational bonds developed towards the crowdsourcer and intermediary, leading to diverse behavioral value-contributions, which exceeded repeat interactions. The case also illustrated that the underlying process of engagement included the emotional response to specific stimuli, which led to delight, involvement, empowerment and/or trust, fostering affective commitment and (planned) indirect value contributions (i.e., WOM, referral, further knowledge contributions, observations, consumption activities). Next to those illustrations of the theoretical concept, the case helped to extend and refine knowledge concerning the underlying mechanisms of the engagement process. First of all, it could be shown that engagement developed differently towards the crowdsourcer and intermediary throughout the process. Hence, participants were able to differentiate stimuli-related experiences and draw separate conclusions. It further showed that stimuli evaluations may depend on prior familiarity with the engagement object and its interactive character. Those rather administrative stimuli, appearing in the pre- and post-participation phase, which were quite familiar for most crowdsourcees, led to mostly satisfaction and trust. In comparison, those rather interactive stimuli in the participation phase, which differed from project to project (i.e., new types of tasks, other test objects), fostered delight, involvement and empowerment, if designed properly (Table 1).

This also relates to the different roles of crowdsourcees, influencing the perception of stimuli and its impact on engagement. It was discussed that from an IS-perspective platform quality and characteristics of the test object may play a role for engagement [26, 27]. The case illustrated that the crowdsourcing platform arose no emotional responses. This might be due to its transactional character and consistency throughout interactions. Here, the goal should be to rather strive for satisfaction and potentially enhance trust into the technology in the long term. The test object however, due to its hedonistic character, led to several emotional responses and arose delight and involvement, leading to longer interaction times and even the desire for further knowledge contributions and observations after participation. It seems to be an important factor that potentially drives affective commitment. Furthermore, from an OB-perspective it was assumed that the task and reward may stimulate emotional responses [28]. In this case, the monetary reward had rather a utilitarian purpose. However, perceived fairness and reliability regarding the transaction process fostered trust over repeat interactions. Moreover, the tasks and test scenarios stimulated delight and involvement due to perceived fun and challenge, which even fostered task over-fulfillment and a more intense interaction on the crowdsourcing platform. Besides, some crowdsourcees mentioned to enjoy having impact on the test object. Thus, perceived relevance of the test object and task may enhance sensed empowerment and eventually affective commitment, as it is predicted by the theory of psychological ownership [32]. From a marketing perspective, the case showed that even (planned) consumption activities could be stimulated due to positive experiences with the task and test object, fostering a positive attitude towards the crowdsourcer and its products. Consequently, from a managerial perspective, it would be effective to design stimuli that foster satisfaction and trust in the pre- and post-participation phase; and delight, involvement and empowerment in the participation phase to enhance engagement.

Nevertheless, those empirical observations are not sufficient to prove relationships, as a single crowdsourcing case was assessed with a limited number of interviews. Yet, the illustrative case can be seen as a pilot study, suggesting a promising methodology and valuable first insights. For future research it is recommended to conduct multiple case studies, including different types of crowdsourcing to identify more engagement-driving mechanisms, patterns, and related stimuli characteristics from a process perspective (e.g., collaborative vs. non-collaborative, paid vs. unpaid, complex vs. micro-tasks, etc.). Additionally, to verify relationships with quantitative research, a survey approach may be applied, which tests for arising drivers and its impact on affective commitment and behavior. Pre- and post-participating engagement states may be compared to verify effects. Besides, experiments with manipulated stimuli may be used to explain concrete effects on engagement outcomes. The provided framework in Figure 4, can serve as a base for future research.

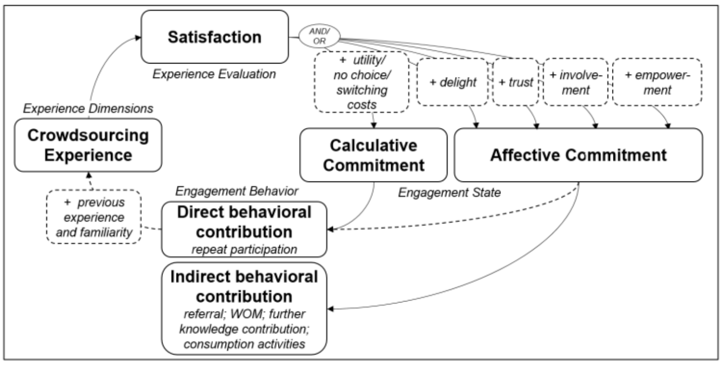


Figure 4. Research Model of the Crowdsourcee's Engagement Process

6 Conclusion

Applying the engagement concept and process to the case of crowdsourcing and deploying an adapted form of SILT as a unique measuring approach is a first step in offering researchers an experience-based perspective on crowdsourcing. The integration of those rather new research fields has the advantage that valuable knowledge for both can be derived. Crowdsourcing facilitates the connectivity of people, organizations and societies via a technological platform. In the center of this research is the IT-mediated *Crowdsourcing Experience*, generated through experience-driving stimuli. Hence, this research contributes to the IS literature, delivering insights on the so far under-researched concept of IT-enabled engagement processes between individuals and entities, from a psychological and behavioral perspective. Additionally, the concept of engagement is considered as a new perspective in relationship marketing research. By illustrating the engagement process with a first case, the aim is to support the progress of the engagement concept from an emergent theme to a more mature construct. Nevertheless, developing a better understanding of the currently realized *Crowdsourcing Experience* and the underlying mechanisms of the engagement process may help practitioners to improve the interaction process and identify engagement opportunities.

Appendix

Findings from the Interview-Assessment

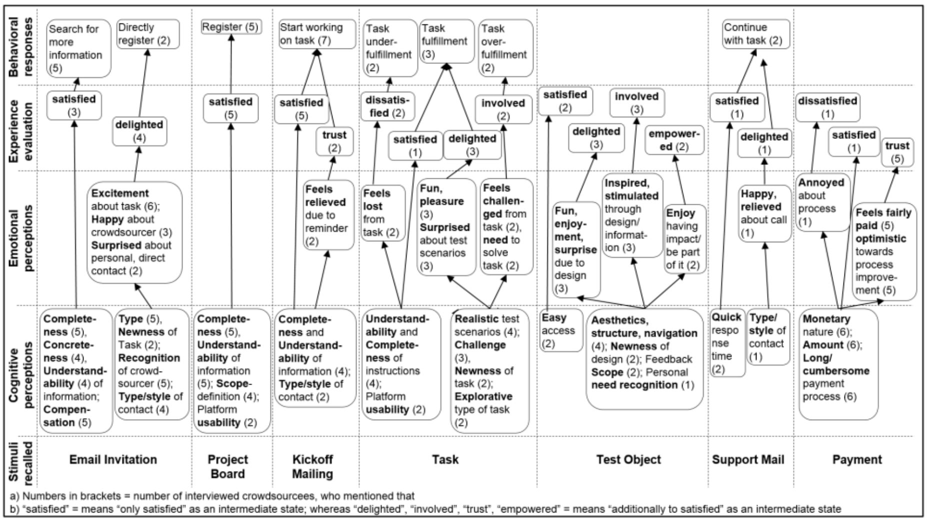


Figure A1. Crowdsourcing Experience Analysis based on SILT-Approach

References

1. Owyang, J.: The state of crowdsourcing in 2015. Crowd Companies™ (2015)
2. Blohm, I., Leimeister, J.M., Krcmar, H.: Crowdsourcing: How to Benefit from (Too) Many Great Ideas. MIS Quarterly Executive 12, 199-211 (2013)
3. Estellés-Arolas, E., González-Ladrón-De-Guevara, F.: Towards an integrated crowdsourcing definition. Journal of Information Science 38, 189-200 (2012)
4. Vuković, M.: Crowdsourcing for enterprises. In: Services-I, 2009 World Conference on, pp. 686-692. IEEE, (2009)
5. Prahalad, C.K., Ramaswamy, V.: Co-creation experiences: The next practice in value creation. Journal of interactive marketing 18, 5-14 (2004)
6. Gassmann, O., Winterhalter, S., Wecht, C.H.: Crowdsourcing: Tipps, damit es gelingt. IO Management 44-50 (2013)
7. Füller, J., Mühlbacher, H., Matzler, K., Jawecki, G.: Consumer empowerment through internet-based co-creation. Journal of Management Information Systems 26, 71-102 (2009)
8. Pedersen, J., Kocsis, D., Tripathi, A., Tarrell, A., Weerakoon, A., Tahmasbi, N., Xiong, J., Deng, W., Oh, O., De Vreede, G.-J.: Conceptual foundations of crowdsourcing: A review of IS research. In: System Sciences (HICSS), 2013 46th Hawaii International Conference on, pp. 579-588. IEEE, (2013)
9. Zogaj, S., Bretschneider, U., Leimeister, J.M.: Managing crowdsourced software testing: a case study based insight on the challenges of a crowdsourcing intermediary. Journal of Business Economics 84, 375-405 (2014)
10. Ye, H.J., Kankanhalli, A.: Investigating the antecedents of organizational task crowdsourcing. Information & Management 52, 98-110 (2015)
11. De Vreede, T., Nguyen, C., De Vreede, G.-J., Boughzala, I., Oh, O., Reiter-Palmon, R.: A theoretical model of user engagement in crowdsourcing. Collaboration and Technology, pp. 94-109. Springer (2013)
12. Sun, Y., Fang, Y., Lim, K.H.: Understanding sustained participation in transactional virtual communities. Decision Support Systems 53, 12-22 (2012)
13. Nguyen, C., Tahmasbi, N., de Vreede, T., de Vreede, G.-J., Oh, O., Reiter-Palmon, R.: Participant Engagement in Community Crowdsourcing. Paper accepted to be presented at the European Conference of Information Systems, Münster, Germany, (2015)
14. Riedl, C., Blohm, I., Leimeister, J.M., Krcmar, H.: The effect of rating scales on decision quality and user attitudes in online innovation communities. International Journal of Electronic Commerce 17, 7-36 (2013)
15. Schulten, M.B., Schaefer, F.: Affective commitment and customer loyalty in crowdsourcing: antecedents, interdependencies, and practical implications. International Review of Retail, Distribution & Consumer Research 25, 516-528 (2015)
16. Brodie, R.J., Hollebeek, L.D., Juric, B., Ilic, A.: Customer engagement: conceptual domain, fundamental propositions, and implications for research. Journal of Service Research 1094670511411703 (2011)
17. Brodie, R.J., Ilic, A., Juric, B., Hollebeek, L.: Consumer engagement in a virtual brand community: An exploratory analysis. Journal of Business Research 66, 105-114 (2013)
18. Sashi, C.: Customer engagement, buyer-seller relationships, and social media. Management decision 50, 253-272 (2012)

19. Bowden, J.L.-H.: The process of customer engagement: A conceptual framework. Journal of Marketing Theory and Practice 17, 63-74 (2009)
20. Mollen, A., Wilson, H.: Engagement, telepresence and interactivity in online consumer experience: Reconciling scholastic and managerial perspectives. Journal of business research 63, 919-925 (2010)
21. Oliver, R.L.: Effect of expectation and disconfirmation on postexposure product evaluations: An alternative interpretation. Journal of applied psychology 62, 480 (1977)
22. Oliva, T.A., Oliver, R.L., Bearden, W.O.: The relationships among consumer satisfaction, involvement, and product performance: A catastrophe theory application. Behavioral Science 40, 104-132 (1995)
23. Rust, R.T., Oliver, R.L.: Should we delight the customer? Journal of the Academy of Marketing Science 28, 86-94 (2000)
24. Kumar, V., Aksoy, L., Donkers, B., Venkatesan, R., Wiesel, T., Tillmanns, S.: Undervalued or overvalued customers: capturing total customer engagement value. Journal of Service Research 13, 297-310 (2010)
25. Vivek, S.D., Beatty, S.E., Morgan, R.M.: Customer engagement: Exploring customer relationships beyond purchase. Journal of Marketing Theory and Practice 20, 122-146 (2012)
26. Attfield, S., Kazai, G., Lalmas, M., Piwowarski, B.: Towards a science of user engagement (position paper). In: WSDM Workshop on User Modelling for Web Applications. (2011)
27. O'Brien, H.L., Toms, E.G.: What is user engagement? A conceptual framework for defining user engagement with technology. Journal of the American Society for Information Science and Technology 59, 938-955 (2008)
28. Macey, W.H., Schneider, B.: The meaning of employee engagement. Industrial and organizational Psychology 1, 3-30 (2008)
29. Algesheimer, R., Dholakia, U.M., Herrmann, A.: The social influence of brand community: Evidence from European car clubs. Journal of marketing 69, 19-34 (2005)
30. Ulrich, D.: Tie the corporate knot: Gaining complete customer commitment. MIT Sloan Management Review 30, 19 (1989)
31. Fuchs, C., Prandelli, E., Schreier, M.: The psychological effects of empowerment strategies on consumers' product demand. Journal of Marketing 74, 65-79 (2010)
32. Spreitzer, G.M.: Psychological empowerment in the workplace: Dimensions, measurement, and validation. Academy of management Journal 38, 1442-1465 (1995)
33. Leonardi, P.M.: When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies. MIS quarterly 35, 147-167 (2011)
34. Jüttner, U., Schaffner, D., Windler, K., Maklan, S.: Customer service experiences: Developing and applying a sequentialincident laddering technique. European Journal of Marketing 47, 738-769 (2013)
35. Koenig-Lewis, N., Palmer, A.: Experiential values over time–a comparison of measures of satisfaction and emotion. Journal of marketing management 24, 69-85 (2008)
36. Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. Weinheim Basel BELTZ, 2015 12. überarbeitete Auflage (2015)
37. Gläser, J., Laudel, G.: Experteninterviews und qualitative Inhaltsanalyse. als Instrumente rekonstruierender Untersuchungen. Wiesbaden VS Verlag für Sozialwissenschaften 2010 4. Aufl. (2010)

Flow in Information Systems Research: Review, Integrative Theoretical Framework, and Future Directions

Raphael Rissler[1,2], Mario Nadj[1,2,3], Marc T. P. Adam[4]

[1] SAP SE, Walldorf, Germany
{raphael.rissler, mario.nadj}@sap.com
[2] Karlsruhe Institute of Technology (KIT), IISM, Karlsruhe, Germany
{raphael.rissler, mario.nadj}@partner.kit.edu
[3] The University of Mannheim, Institute for Enterprise Systems, Mannheim, Germany
nadj@es.uni-mannheim.de
[4] The University of Newcastle, School of Design Communication and IT, Newcastle, Australia
marc.adam@newcastle.edu.au

Abstract. As information systems (IS) are increasingly able to create highly engaging and interactive experiences, the phenomenon of *flow* is considered a promising vehicle to understand pre-adoptive and post-adoptive IS user behavior. However, despite a strong interest of researchers and practitioners in flow, the reliability, validity, hypothesized relationships, and measurement of flow constructs in current IS literature remain challenging. By reviewing extant literature in top IS outlets, this paper develops an integrative theoretical framework of flow antecedents, flow constructs, and flow consequences within IS research. In doing so, we identify and discuss four major flow streams in IS research and indicate future research directions.

Keywords: Flow theory, flow measurement, flow streams, human-computer interaction, integrative theoretical framework

1 Introduction

In today's digital economy, information systems (IS) are both, a significant investment for companies and an integral part of employees' daily work [1]. Due to technological developments, such as multi-media-rich user interfaces (UIs), IS are able to create highly engaging and interactive experiences [1]. More specifically, the design and implementation of IS plays an important role in whether or not users have holistic experiences such as "flow" when interacting with technology. Moreover, in today's technology landscape, most work-related tasks are at least to some extent IT-mediated. Hence, studying how flow affects pre-adoptive as well as post-adoptive IS use has been acknowledged of theoretical and practical significance [1–3]. Thereby, flow is adopted from the reference discipline of psychology and refers to "the holistic sensation that people feel when they act with total involvement" [4, p. 36].

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Rissler, R.; Nadj, M.; Adam, M.T.P. (2017): Flow in Information Systems Research: Review, Integrative Theoretical Framework, and Future Directions, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1051-1065

However, despite the intense usage of flow-related constructs within IS research, its reliability and validity still remains low [5–9]. Novak and colleagues identified 13 different flow constructs with an average usage rate of only four per study [5]. Conceptually, the inconsistencies also concern the hypothesized relationship between flow and other constructs [8]. Finneran and Zhang [7] concluded that the "diverse flow models demonstrate the different understandings of antecedents, flow experiences, and consequences" [7, p. 98]. Moreover, most flow constructs used by IS researchers only partially overlap with the constructs and measurements suggested by the reference discipline of psychology [4, 10, 11]. In summary, it can be concluded that current IS research summarizes the usage of flow within its discipline as "too broad and ill-defined due to the numerous ways it has been operationalized, tested, and applied." [12, p. 227]

In this paper, we review 43 articles in top IS outlets pursuing the following research question: *What is the state-of-the-art in flow research within top IS outlets?* Our SLR builds upon existing reviews [6, 8, 13] and extends these studies in several ways. We complement the literature-based discussions on flow by Finneran and Zhang [6] and Siekpe [13] with a structured approach including detailed information about the search approach, used databases, search strategy, and study selection criteria. Specifically, we expand the work by Finneran and Zhang [6] by flow dimensions and consequences, as well as incorporate the six flow constructs proposed by Siekpe [13]. Further, building on the nine stream suggestions by Mahnke et al. [8], we consolidate four streams of flow literature by analyzing the operationalization of the identified constructs. In addition, on the basis of our SLR and previous work [6, 8, 13], we synthesize the knowledge of flow in those four streams and develop an integrative theoretical framework, consisting of overarching flow categories, as well as sub-categories (cf. Figure 3). This framework can serve as a 'route map' in understanding the relations between various flow components and their interactions, as well as provide different academic perspectives on flow [14]. Finally, we cluster the identified articles in this framework accordingly to illustrate the most prominent streams and gaps (cf. Table 2).

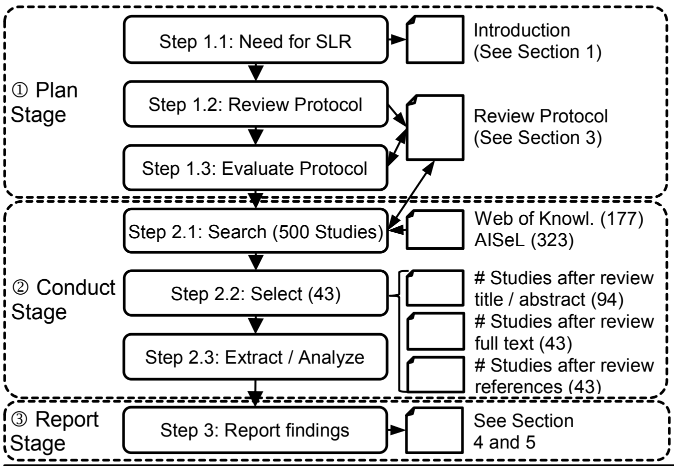
This paper makes five key contributions to IS research and practice. First, we introduce a comprehensive, integrative theoretical framework of flow in IS research [14]. This high-level framework can support both, IS researchers and practitioners to conceptualize the flow phenomenon and guide the design of IS artefacts. Second, we provide a detailed overview of four major flow streams in IS research and position these streams within our integrative theoretical framework. Third, we summarize the results of the literature with respect to the major antecedents and consequences of flow. Fourth, we provide a detailed overview on the commonly used flow constructs within IS research and identify the major challenges in their operationalization. Finally, our review provides suggestions for further research within IS.

2 Fundamentals of Flow

Flow was first investigated by psychologist Mihaly Csikszentmihalyi, who developed a theory of flow in the 1970s based on qualitative interviews with individuals performing (autotelic) activities in a non-professional context without extrinsic rewards

[4]. During the analysis a pattern emerged in which individuals were fully immersed and concentrated on a task at hand – the so called "flow experience" [4]. Further, Csikszentmihalyi differentiated between different degrees of flow ranging from micro-flow (e.g., perceived at taking a coffee break) and deep-flow (e.g., perceived while painting a picture) [4]. Initially, researchers hypothesized that flow experience occurs only in cases, where the performer of an activity does not receive any extrinsic reward (e.g., financial benefits) [4]. However, further studies showed that flow constitutes a general phenomenon occurring in both, extrinsically (e.g., working environment) [2] and intrinsically (e.g., painting, music) motivated activities [4]. Due to this high generalizability and recent enhancements in IS capabilities to foster flow (e.g., via providing multi-media-rich task support), the concept of flow has been widely adopted by researchers to understand user behavior in engaging and interactive technology contexts [1]. For instance, based on an analysis of 43 employees using an e-mail application, researchers established a direct link between flow experience and actual technology use. In the same study, researchers also found a correlation between flow and other constructs, such as perceived communication quality and perceived effectiveness [2]. Generally, the flow phenomenon has been applied in various domains, such as E-Learning [15], E-Commerce [16], Web-Sites [17], Games [18], and Virtual Worlds [19] in order to explain and study user perception and behaviors. Building upon several empirical flow studies in computer-mediated environments, scholars examined the concept through three different angles: (1) *flow antecedents*, (2) *flow experience*, and (3) *flow consequences* [3, 6, 12, 20]. However, the conceptualizations, antecedents, and consequences vary across studies (e.g., [9, 21]).

3 Method



Distribution of Flow Articles in Top IS Outlets

Source	# Hits	# Selected	%
MISQ	28	6	14
ISR	27	1	2.3
JMIS	26	3	7
EJIS	12	2	4.7
I&M	49	16	37.2
ISJ	12	2	4.7
JAIS	13	2	4.7
JSIS	10	0	0
ICIS	223	8	18.6
ECIS	100	3	7
Sum	500	43	~ 100

MISQ: MIS Quarterly; **ISR**: Information Systems Research; **JMIS**: Journal of Management Information Systems; **EJIS**: European Journal of Information Systems; **I&M**: Information & Management; **ISJ**: Information Systems Journal; **JAIS**: Journal of the AIS; **JSIS**: Journal of Strategic Information Systems; **ICIS**: International Conference on Inform. Systems; **ECIS**: European Conference on Inform. Systems

Figure 1. Stages of the SLR and Distribution of Flow Articles

In order to evaluate the current state of flow research within top IS outlets, we conduct a SLR following the guidelines by Kitchenham [22]. The SLR is subdivided into three stages (plan, conduct, and report; cf. Figure 1). During the **plan stage**, we identified the need for a SLR. In a second step, we developed a review protocol and evaluated it. During the **conduct stage**, we executed the search, selected appropriate studies, and analyzed them. Finally, during the **report stage**, we documented our findings.

Research Questions. To keep our systematic review focused, and to answer the overarching research question, we defined several subordinate questions (cf. Table 1).

Table 1. Subordinate Research Questions

RQ #	Research Questions
RQ1	What are the different streams of flow in top IS outlets and how can they be conceptualized into an integrative theoretical framework?
RQ2	What are the antecedents of flow in top IS outlets?
RQ3	How is the flow phenomenon conceptualized and operationalized in top IS outlets?
RQ4	What are the consequences of the flow phenomenon in top IS outlets?

Search Strategy. To support the search process (Step 2.1, Figure 1), we first selected libraries based on our research questions. As our goal was to provide a holistic overview on the state-of-the-art in flow research within the IS domain, our 'field' is the discipline of IS. To get an overview of high quality studies, we decided to include top-tier **IS journals** (cf. Figure 1) from the *IS Senior Scholars' basket of eight*. We also decided to include two major **IS conferences**, namely the *International Conference on Information Systems* (ICIS) and the *European Conference on Information Systems* (ECIS). The outlets were carefully selected on the basis of a ranking list (http://www.core.edu.au/conference-portal), as well as suggestions made by literature to include especially journals and conferences with high quality and reputation [23]. However, it should be mentioned, that we did not include research-in-progress papers. Based on the identified outlets and field of interest, we selected the *ISI Web of Knowledge* as **database to search** for the IS journals. In addition, we selected the *AIS Electronic Library* (AISeL) to retrieve conference proceedings (ICIS, ECIS). The **search string** to conduct our systematic search (Step 2.1, Figure 1) was developed in several steps. First, we extracted "flow" as a starting term from our research questions. Second, we used the term "flow AND information systems" to search for publications within IS using Google Scholar. By reviewing the first 20 hits and by sorting out papers without a focus on the psychological phenomenon of flow, we identified two highly cited papers. Namely, Agarwal and Karahanna [1] and Hsu and Lu [24]. By reviewing the full text, we extracted the term "cognitive absorption" and "cognitive engagement" as highly relevant flow derivations. In a third step, we searched for synonyms but did not find any appropriate synonyms for our study context. Finally, we used Boolean-operators in order to create the final search string: *flow OR cognitive engagement OR cognitive absorption*. Next, we applied the final search string to the **title**, **abstract,** and **keywords section** of publications in the specified digital libraries. We did **not limit** our search to a specific **time period**, as the aim of our SLR was to provide a holistic overview. The overall hits, as well as the final number of selected studies and the percentage distribution are depicted in Figure 1.

Study Selection Criteria. We carefully defined the following inclusion and exclusion criteria: (1) Only empirical studies were included, (2) studies using flow, cognitive engagement, or cognitive absorption in their hypothesis development were included, whereas (3) studies not referencing to the psychological phenomenon of flow were excluded. In the publication selection process (Step 2.2, Figure 1) the criteria were applied to title, abstract, and keyword section excluding 406 inappropriate studies. Second the criteria were applied to full text, again excluding 51 studies. Finally, we reviewed the references but did not find any additional publications, as our SLR is focused on the defined IS outlets and we already found all studies in the selected databases. In summary, we found 43 relevant studies.

4 Flow Streams within IS Research

As depicted in Figure 2, our SLR reveals that starting from 2002, flow received wider attention in IS research and is still very active with an increasing amount of publications. Thereby, on average three papers were published per year in top IS outlets from 2002 to 2016. As 2016 is ongoing, only two studies were found in this year. In the next sections we describe the identified four streams of flow research in IS.

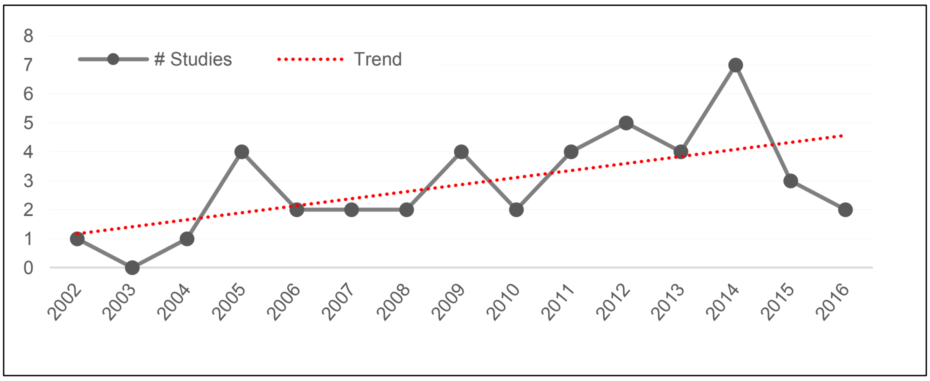


Figure 2. Distribution of Studies over the Years

Stream 1 – Jackson / Marsh / Ghani / Deshpande / Supnick / Rooney. The first stream [3, 20, 25] is based on the work of Jackson and Marsh [25] as well as Ghani et al. [3, 20] entailing nine references. Both research groups use concentration as flow construct consisting of identical items, such as "My attention was focused entirely on what I was doing." [25, p. 34] and "Attention is focused on activity" [20, p. 390]. In addition, the construct of "autotelic experience" used by Jackson and Marsh [25] and the construct of "enjoyment" used by Ghani et al. [3, 20] are highly related as both entail items of positive emotions such as fun or enjoyment [3, 20, 25]. Further, articles in this stream cite and refer particularly to both research groups when conceptualizing flow. For instance, Guo et al. [26] use concentration according to Ghani et al. [20] and combine it with dimensions from Jackson and Marsh [25]. However, it is important to note that some studies in this stream only refer to one research group (primarily Ghani et al. [3]). We assume the reason behind this dominance resides in Ghani et al.'s [3, 20] specific IS focus, whereas Jackson and Marsh [25] are originally allocated in the sports domain. Jackson and Marsh [25] developed the flow state scale in accordance with the proposed characteristics of flow suggested by Csikszentmihalyi [4, 11]. Thereby, the researchers used the following dimensions: (1) challenge-skill balance, (2) clear goals, (3) unambiguous feedback, (4) autotelic experience (5) action-awareness merging, (6) sense of control, (7) loss of self-consciousness, (8) transformation of time, and (9) concentration on the task at hand. Among these dimensions, empirical evidence has shown that **challenge-skill balance**, **clear goals**, and **unambiguous feedback** represent major antecedents of flow [27]. Supported by theory [11] and evidence in literature [28–30], the construct of **autotelic experience** (e.g., operationalized by authors as enjoyment or positive affect [8, 26]) constitutes an important outcome variable of flow. The other characteristics pertain to the phenomenon itself. Particularly, **action-awareness merging** refers to a state, where due to a deep level of involvement, an activity becomes automatic or spontaneous [25]. **Sense of control** refers to the feelings and perceptions of being in charge of the interaction [25]. **Loss of self-consciousness** is described as disappearance of concerns for the self [25]. Further, the **transformation of time** construct proposed by Jackson and Marsh [25] emphasized the altering of time (e.g., slower or faster) [25]. Finally, **concentration on the task at hand** refers to feelings of being focused and concentrated on the task at hand [25]. Ghani et al. [3, 20] conceptualize flow with two main characteristics. Enjoyment resulting from the activity and total concentration [3, 20]. This stream (as depicted in Table 2) predominantly investigates flow in the context of E-Learning and E-Commerce. Thereby flow is investigated multi-dimensionally using constructs such as concentration (66.7 %), and sense of control (55.6 %). With regard to the antecedents, artefact-related antecedents (88.9 %) are highly used. As for flow consequences, this stream predominantly uses behavior-related constructs (77.8 %).

Stream 2 – Agarwal / Karahanna / Skadberg / Kimmel. Within the second stream [1, 31] (17 reference articles), the most common conceptualization of flow is mainly based on Agarwal and Karahanna [1]. The concept of cognitive absorption (CA) includes five dimensions: (1) curiosity, (2) control, (3) focused immersion, (4) temporal dissociation, and (5) heightened enjoyment. Thereby, **control (control of interaction)** is defined as "the user's perception of being in charge of the interaction" [1, p. 673] and **curiosity** refers to "heightened arousal of sensory and cognitive curiosity" [1, p. 668] (cf. third stream). The dimension of **focused immersion** "suggests that all of the attentional resources of an individual are focused on the particular task, thereby reducing the level of cognitive burden associated with task performance." [1, p. 675] **Temporal dissociation** is defined as "the inability to register the passage of time while engaged in interaction" [1, p. 673] and finally, **heightened enjoyment** is "capturing the pleasurable aspects of the interaction" [1, p. 673]. In the same veins like Agarwal and Karahanna [1], Skadberg and Kimmel [31] also conceptualized flow with the dimensions of **enjoyment** and **time distortion** [31]. Thereby, time distortion as well as enjoyment are highly related to the construct of temporal dissociation and enjoyment suggested by Agarwal and Karahanna [1]. This stream (as depicted in Table 2) predominantly investigates flow in the context of the Web, followed by E-Learning and Virtual Worlds. Thereby, most authors in this stream investigate flow as a second-order multidimensional phenomenon through the lens of CA [1]. Constructs of focused immersion (94.1 %) and transformation of time (94.1 %) are used most dominantly. Further, artefact-related antecedents (41.2 %) are dominant, followed by person-related antecedents (23.5 %). As for the flow consequences in this stream, behavior-related constructs (82.4 %) are prior to cognition-related consequences (64.7 %).

Stream 3 – Webster / Trevino / Ryan / Ho. The third stream [2, 32, 33] entails the lowest number of references (seven articles). It originates from the studies of Webster et al. [2] and Trevino and Webster [33] who both suggested four dimensions of flow experience: (1) sense of control over the interaction, (2) curiosity, (3) intrinsic interest, and (4) attention focus [2]. In a later study, Webster and Ho [32] conceptualize the last three dimensions as cognitive engagement. Similar to the definition in the first stream, the dimensions of **sense of control** refers to the feelings of control, as well as the actual control over the interaction [2], whereas **curiosity** illustrates the arousal of sensory or cognitive curiosity [2] (cf. second stream). **Intrinsic interest** is defined as cognitive arousal as well as imagination [2]. The construct of **attention focus** suggests, that the "attention is narrowed to a limited stimulus field, filtering out irrelevant thoughts and perceptions." [2, p. 413] In this stream, authors predominantly investigate the phenomenon of flow through the lens of cognitive engagement [32]. As depicted in Table 2, authors within this stream primarily use the constructs of attention focus and intrinsic interest (both 100 %), followed by curiosity (71.4 %). With regard to the antecedents, artefact-related antecedents (71.42 %) are used most widely. Further, behavior-related constructs (85.7 %) are dominant.

Stream 4 – Novak / Hoffman / Yung / Engeser / Rheinberg. The fourth stream [5, 10, 34] addresses flow from a wider and more general perspective. Hence in Figure 3 this stream is positioned at a higher conceptual level than the other three streams. Novak et al. [5] and Hoffman and Novak [34] contributed to this stream (ten articles) by measuring flow as a one-dimensional construct with a narrative description of flow experience [5, 34]. The operationalization consists of items such as "In general, how frequently would you say you have experienced "flow" when you use the Web?" [5, p. 28]. Similarly, Engeser and Rheinberg [10] also contributed to this stream on a higher conceptual level as they propose a comprehensive flow short scale to measure flow during all activities [10]. Conducting a factor analysis, the researchers found two overreaching and broadly defined factors which they labeled "fluency" and "absorption" consisting of items such as "My thoughts/activities run fluid and smoothly", or "I am completely lost in thoughts" [10, p. 170]. This stream predominantly investigates flow in the context of the Web and Games. With regard to the antecedents, artefact-related antecedents (50 %) are used most. As for flow consequences, all studies include behavior-related constructs in this stream.

Table 2. Flow Streams in Top IS Outlets

			Antecedents							Absorption						CNT		CE			Conse.			
	Authors	Stream Ref.	Pers.(P)	P & A	Artef.(A)	A & T	Task(T)	T & P	Flow	AAM	FIM	CON	TDI	TRT	LSC	COI	SOC	CUR	ATF	INI	Cognit.	Affect.	Behav.	Perform.
Stream 1	Guo et al. [26]	[20, 25]		X			X	X				X		X	X						X	X	I	
	Zhang et al. [35]	[20]		X	X							X									X	X	I	
	Guo et al. [36]	[20, 25]		X	X		X	X				X		X	X		X				X	X	I	P
	Nah et al. [30]	[3]		X	X																X		I	
	Nah et al. [37]	[3]		X	X																X		I	
	Guo & Poole [27]	[25]			X		X	X		X		X	X		X		X							
	Phang et al. [38]	[3]	X	X	X							X					X							P
	Kamis et al. [39]	[3]			X		X										X						I	
	Koufaris [40]	[3]			X			X				X					X						I	
Stream 2	Visinescu et al. [41]	[1]			X						X		X					X			X		I	
	Mimoun et al. [42]	[1]									X		X			X		X						O
	Wang & Hsu [43]	[31]					X	X			X		X								X			O
	Lowry et al. [44]	[1]			X						X		X			X		X					I	
	Goel et al. [45]	[1]	X	X	X						X		X					X			X		A	P
	Goel et al. [46]	[1]		X							X		X										I	
	Lee et al. [47]	[1]			X						X		X								X	X	I	
	Chandra et al. [19]	[1]	X			X					X		X			X		X			X		I	
	Weniger et al. [48]	[1]	X		X						X		X			X		X			X	E	I	
	Goel et al. [49]	[1]		X				X					X										I	
	Deng et al. [50]	[1]									X		X			X		X			X	X	I	
	Xue & Hock-Hai [18]	[1]			X						X		X			X		X					I	
	Jia et al. [51]	[1]									X		X			X		X				X	A	
	Wakefield et al. [29]	[1]									X		X			X		X			X	E	I	
	Shang et al. [52]	[1]									X		X			X		X			X		A	
	Saadé & Bahli [53]	[1]									X		X								X		I	
	Hess et al. [54]	[1]	X		X						X										X			O
Stream 3	Yi et al. [16]	[32]			X														X	X			I	
	Zhang et al. [55]	[2]		X	X													X	X	X			I	
	Moon et al. [17]	[2, 33]	X													X		X	X	X			A	
	Animesh et al. [56]	[2]		X	X													X	X	X			I	
	Scott & Walczak [57]	[32]																	X	X	X		I	
	Webster & Ahuja [58]	[32]		X	X													X	X	X			I	O
	Jiang & Benbasat [21]	[2]			X											X		X	X	X				
Stream 4	Rodríguez et al. [15]	[5]	X	X	X				X														❶	
	Bilgihan et al. [59]	[5]			X				X												X		I	
	Mahnke et al. [8]	[10]	X				X	X	X													X	I	
	Huang et al. [60]	[5]		X					X												X	X	I	
	Kim et al. [28]	[5]		X	X				X														I	
	Sharkey et al. [61]	[3, 5]							X														I	P
	Mahnke et al. [9]	[10]							X													X	I	
	Theotokis et al. [62]	[5]										X				X		X			X	X	I	
	Ha et al. [63]	[5]		X	X				X														I	
	Hsu & Lu [24]	[5]			X				X														I	

AAM: Action Awareness Merge; **FIM**: Focused Immersion; **CON**: Concentration; **ATF**: Attention Focus; **TDI**: Temporal Dissociation; **TRT**: Transformation of Time; **LSC**: Loss of Self-Consciousness; **COI**: Control of Interaction; **SOC**: Sense of Control; **CUR** Curiosity; **INI**: Intrinsic Interest; **P**: Perceived; **I**: Intention; **O**: Objectively; **A**: Actual Behavior; **E**: Enjoyment; **CNT**: Control; **CE**: Cognitive Engagement; **❶**: Intention & Actual Behavior

5 Integrative Theoretical Framework of Flow in IS Research

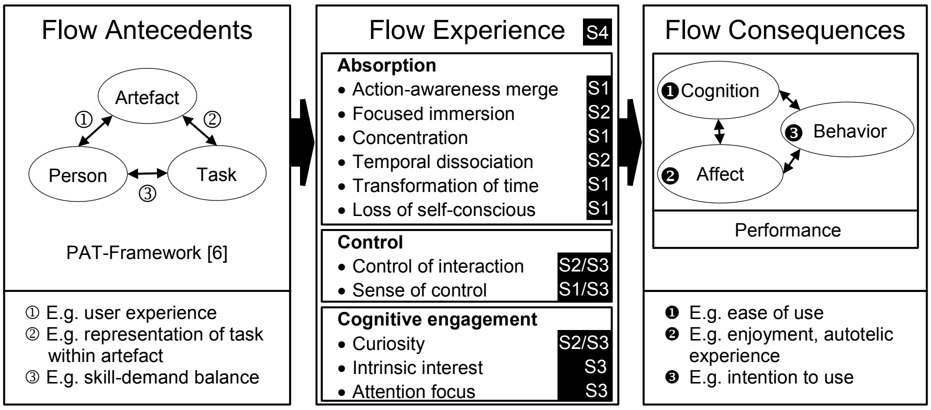
In order to address RQ1, we follow the approach of Baumeister and Leary [14] and conceptualize the results of the extant literature in an integrative theoretical framework [14]. By reviewing our final set of primary studies, several interesting patterns appear. First, all of the studies explicitly or implicitly subdivide their research models into *flow antecedents*, *flow experience*, and *flow consequences* [26]. As the separation of flow into this threefold pattern is also generally agreed upon in IS literature [3, 6, 12, 20], we adopt this separation for our theoretical framework.

Flow antecedents. We found that some of the antecedents are related to **tasks-characteristics,** such as clarity of goals [36], and immediate feedback on the task [26]. Other antecedents of flow are related to the **UI**. For instance, in one study, scholars manipulate filler interfaces on travel booking sites to evaluate the effects on flow and perceived waiting time [47]. In a third category of studies, researchers investigate **person-characteristics**, such as gender [54]. Therefore, we follow the approach of Finneran and Zhang [6] and sub-divide antecedents of flow into (1) **P**erson, (2) **A**rtefact, and (3) **T**ask (*PAT-Framework)*. Moreover, in reviewing the results of the extant literature, we do find several constructs rather allocated between the categories of person, artefact, and task. For instance, researchers investigated the influence of skill-demand balance on flow [8, 27, 43]. As such, the balance between demand and skill is neither a pure task-characteristic nor a pure person-characteristic. Similarly, other characteristics, such as user experience (allocated between person and artefact) and the representation of tasks (allocated between artefact and task) are allocated between two characteristics. Thus, we introduce categories between the defined antecedents-sections of person, task, and artefact as depicted in Figure 3.

Flow consequences. We identified three major categories of flow consequences: (1) **cognition**, (2) **affect**, and (3) **behavior**. The proposed categories are well-rooted in recent social psychology research, suggesting that attitude consists of these three distinct dimensions [64]. The affect-related component constitutes the hedonic aspect of the attitude towards an IS, such as moods and emotions. In contrast, the cognition-related component constitutes the utilitarian aspect based on beliefs such as ease of use or usefulness. The behavior-related component depicts the response resulting from affect and cognition (e.g., the intention to use an IS). As we found studies investigating the flow impact on performance, we also add performance to the consequence-section.

Flow experience. With regard to flow, we do not consider autotelic experience as part of flow because Csikszentmihalyi [4, 11] argued that positive affect and flow are two distinct constructs. Thereby, the suggestions made by theorists are also relined by evidence of recent research in this direction [28–30]. In order to integrate the various flow constructs used in IS research (cf. stream section), we first extracted constructs related to the flow dimension of **absorption** [8]. By reviewing six stream research teams [1–3, 20, 25, 32], we extract four constructs. In a second step, we analyze the construct definition and operationalization and identify two distinct constructs, namely, **focused immersion** and **concentration**. Thereby the constructs differ in conceptualization as well as operationalization. **Concentration** consists of items measuring focus, concentration, absorption, and attention [3, 20, 25], whereas **focused**

immersion measures absorption, immersion, the blocking of other attentional demand, and if an individual's attention is distracted easily or not [1]. Focused immersion "suggests that all of the attentional resources of an individual are focused on the particular task, thereby reducing the level of cognitive burden associated with task performance" [1, p. 675]. In a second step, we extract and compare different conceptualizations of time-related constructs used by three authors [1, 25, 31]. As result, we identify two distinct constructs. **Temporal dissociation** emphasize a lost sense of time and a faster time passage [1, 31], whereas **transformation of time** emphasizes the altering of time (slower, faster or stop of time) [25]. As Jackson and Marsh [25] also proposed constructs related to absorption, we add **loss of self-consciousness** and **action-awareness merging** to the integrative theoretical framework. Next, we extract and integrate dimensions used to measure **control**. Thereby, we identify three control-related constructs suggested by literature [1, 2, 25]. By reviewing the conceptualization as well as the operationalization, we identify two distinct constructs. **Control of the interaction** refers to the feelings and perceptions of being in charge of the interaction but also to the actual control over the interaction [1, 2]. In contrast, **sense of control** is solely referring to feelings of being in charge [25]. Finally, by complementing our framework with the dimension of **cognitive engagement** proposed by Webster and Ho [32] and the related dimensions **curiosity**, **intrinsic interest,** and **attention focus** [32], we finalize our integrative theoretical framework as depicted in Figure 3. To provide a comprehensive overview, the framework also lists all research teams in the four streams and maps them to the identified constructs.



Stream	Authors	# Ref
S1	Jackson & Marsh 1996 [25], Ghani 1991 [3], Ghani & Deshpande 1994 [20]	9
S2	Agarwal & Karahanna 2000 [1], Skadberg & Kimmel 2004 [31]	17
S3	Webster et al. 1993 [2], Webster & Ho 1997 [32], Trevino & Webster 1992 [33]	7
S4	Novak et al. 2000 [5], Hoffman & Novak 1996 [34], Engeser & Rheinberg 2008 [10]	10

Figure 3. Integrative Theoretical Framework of Flow in IS Research

6 Discussion, Future Directions and Conclusion

Despite the growing relevance of the flow construct within IS research for understanding user behaviour and informing the design of IS artefacts, challenges remain in terms of its conceptualization, reliability, and validity [5–9]. We provided a holistic overview on the state-of-the-art in flow research within top IS outlets. In the following we discuss our results along the previously defined four research questions.

Discussion. As for **RQ1**, we identify four major flow streams and comprehensively describe them in Section 4. Reviewing the extant literature, we present an integrative theoretical framework conceptualizing the state-of-the-art in flow research (cf. Figure 3). Further, we map the identified stream research teams to our framework identifying overlaps and differences across the streams. Such understanding supports IS researchers and practitioners to get an overview of the major dimensions to consider when investigating the flow phenomenon. As for **RQ2**, pertaining to the antecedents of flow, we find that most studies use artefact-related antecedents (25 studies) and investigate their effect on flow. The second most used antecedent category is located between person and artefact (15 studies). Furthermore, seven studies investigate antecedents located between person and task (e.g., challenge-skill balance) and another seven studies focus on person-related antecedents, such as skills and personal innovativeness. Finally, six studies use task-related antecedents, such as clarity of goals and investigate. Pertaining to the operationalization of flow in IS research (**RQ3**), our review reveals 11 distinct flow constructs. As depicted in Figure 3, these constructs can be categorized into absorption, control, and cognitive engagement. By mapping the four major streams to the different flow constructs, it becomes apparent that there is a different emphasis on flow dimensions across the literature. While some streams focus on a more general understanding of flow (S4), other streams aim at a more detailed understanding of absorption and control (S1), or focus primarily on cognitive engagement (S3). Finally, with respect to the flow consequences (**RQ4**), our review reveals that almost all of the reviewed studies measure behavior-related consequences (37 studies). The operationalization of this measurement varies though, with 32 studies focusing on behavior-related intentions (e.g., continuance intention, or the intention to buy online), while five studies measured actual user behavior (e.g., actual technology usage or actual continuance in an E-Learning course). Further, there is also a focus on cognition and affect-related consequences of flow. In 20 studies, cognition-related consequences of flow (e.g., ease of use) are measured, whereas 12 studies describe affect-related consequences (e.g., enjoyment). With regard to performance, four studies assessed perceived performance as flow consequence. Finally, another four studies measured performance outcomes objectively (e.g., time for task completion).

Future directions. Our findings suggests several important future directions in IS flow research. First, four major streams were identified in this SLR. However, it became apparent that some streams address flow from a wider and more general perspective, while others apply a more detailed perspective, focusing on different dimensions of flow. In order to contribute to a profound understanding of flow in user experience engineering, future research may reconcile the different views and conceptualizations of flow, thereby establishing a unified theory of flow in IS research.

Second, with regard to flow antecedents, the study of Finneran and Zhang [6] proposes testable propositions that future research could address in a series of (controlled) lab experiments. For instance, understanding whether a clear fit between task and artefact leads to a flow experience could comprise a promising starting point in this direction. In addition, future research should also examine the antecedents located between person and task (e.g., challenge-skill balance), user-related antecedents (e.g., skill, gender etc.), and task-related antecedents (e.g., clarity of goals), which are so far scarcely addressed. With regard to the antecedent category of IS artefacts, which represents the category with the highest number of research studies, many undiscovered artefact characteristics are still not investigated using the lens of flow theory. For instance, future research may uncover if and how IT-mediated interruptions influence the perception of flow and what characteristics of the interruption are affecting flow. Third, additional work is needed on flow consequences. Our SLR revealed that objectively measured performance outcomes are scarce (four studies). Thereby, literature reports contradicting findings in whether flow leads to higher performance or not, hinting at the importance of considering further characteristics of the user environment, which might explain such differences. Thus, using objectively measured performance outcomes may provide valuable new insights for this ongoing discussion, particularly when applied to different contexts. Finally, our SLR reveals that at this stage only one study in top IS outlets employed neurophysiological measurements to investigate flow [43]. However, considering neurophysiological measurements of flow seems to be a promising research avenue, as flow may appear only briefly during activities and such neurophysiological measurements enable the researcher to analyze flow in situations without interrupting the user. Hence, future studies may put further emphasis on complementing self-reported flow measurement scales with neurophysiological measurements, such as electroencephalography, eye tracking, skin conductance, and heart rate.

We are aware that our paper has **limitations**. Due to the focus on top IS outlets (basket of eight, ECIS, and ICIS), promising articles from other reference disciplines (e.g., psychology) were sorted out. However, including top IS outlets in SLR represents a common practice within IS research as "major contributions are likely to be in the leading journals" and conferences [23, p. xvi]. Further, any bias in the selected keywords may also provoke a bias in the conceptualization. To reduce this probability, we carefully subdivided our SLR activities into three stages (plan, conduct, and report; cf. Figure 1) following the structured guidelines by Kitchenham [22].

Conclusion. With the advances in user experience engineering, the phenomenon of flow has become increasingly relevant for IS research and practice. Designing highly interactive and engaging interfaces requires a profound understanding of the flow phenomenon and its role in user experience in pre- and post-adoptive scenarios. By identifying the various streams of flow-related studies in IS research and conceptualizing these streams in an integrative theoretical framework along the dimensions of absorption, control, and cognitive engagement, we hope that this paper contributes to reconcile the numerous ways the flow construct has been operationalized in IS research. We believe that a unified conceptualization of flow in future research will be a cornerstone of user experience engineering and the design of engaging IS artefacts.

References

1. Agarwal, R., Karahanna, E.: Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. MIS Q. 24, 665 (2000).
2. Webster, J., Trevino, L.K., Ryan, L.: The dimensionality and correlates of flow in human-computer interactions. Comput. Human Behav. 9, 411–426 (1993).
3. Ghani, J.A., Supnick, R., Rooney, P.: The experience of flow in computer-mediated and in face-to-face groups. ICIS Proc. Paper 9 (1991).
4. Csikszentmihalyi, M.: Beyond Boredom and Anxiety. Jossey-Bass, San Francisco (1975).
5. Novak, T., Hoffman, D., Yung, Y.: Measuring the customer experience in online environments: A structural modeling approach. Mark. Sci. 19, 22–44 (2000).
6. Finneran, C.M., Zhang, P.: A person-artefact-task (PAT) model of flow antecedents in computer-mediated environments. Int. J. Hum. Comput. Stud. 59, 475–496 (2003).
7. Finneran, C.M., Zhang, P.: Flow in computer-mediated environments: Promises and challenges. Commun. AIS. 82–101 (2005).
8. Mahnke, R., Benlian, A., Hess, T.: Flow experience in information systems research: Revisiting its conceptualization, conditions, and effects. ICIS Proc. 1–22 (2014).
9. Mahnke, R., Wagner, T., Benlian, A.: Flow experience on the web: Measurement validation and mixed method survey of flow activities. ECIS Proc. (2012).
10. Engeser, S., Rheinberg, F.: Flow, performance and moderators of challenge-skill balance. Motiv. Emot. 32, 158–172 (2008).
11. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper and Row, New York (1990).
12. Choi, D.H., Kim, J., Kim, S.H.: ERP training with a web-based electronic learning system: The flow theory perspective. Int. J. Hum. Comput. Stud. 65, 223–243 (2007).
13. Siekpe, J.S.: An Examination of the multidimensioinality of flow construct in a computer-mediated environment. J. Electron. Commer. Res. 6, 31–43 (2005).
14. Baumeister, Roy, F., Leary, Mark, R.: Writing narrative literature reviews. Rev. Gen. Psychol. 3, 311–320 (1997).
15. Rodríguez-Ardura, I., Meseguer-Artola, A.: E-learning continuance: The impact of interactivity and the mediating role of imagery, presence and flow. Inf. Manag. 53, 504–516 (2016).
16. Yi, C., Jiang, Z. (Jack), Benbasat, I.: Enticing and engaging consumers via online product presentations. J. Manag. Inf. Syst. 31, 213–242 (2015).
17. Moon, Y.J., Kim, W.G., Armstrong, D.J.: Exploring neuroticism and extraversion in flow and user generated content consumption. Inf. Manag. 51, 347–358 (2014).
18. Xue, Y., Hock-Hai, T.: The conflict between absorption and self-control in playing computer games. ICIS Proc. 132 (2008).
19. Chandra, S., Srivastava, S., Theng, Y.: Cognitive absorption and trust for workplace collaboration in virtual worlds. J. Assoc. Inf. Syst. 13, 797–835 (2012).
20. Ghani, J.A., Deshpande, S.P.: Task characteristics and the experience of optimal flow in human-computer interaction. J. Psychol. 128, 381–391 (1994).
21. Jiang, Z., Benbasat, I.: Virtual product experience. J. Manag. Inf. Syst. 21, 111–147 (2005).
22. Kitchenham, B., Charters, S.: Guidelines for performing systematic literature reviews in software engineering. Engineering. 45, 1051 (2007).

23. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. MIS Q. 26, xiii–xxiii (2002).
24. Hsu, C.-L., Lu, H.-P.: Why do people play on-line games? An extended TAM with social influences and flow experience. Inf. Manag. 41, 853–868 (2004).
25. Jackson, S.A., Marsh, H.W.: Development and validation of a scale to measure optimal experience: The flow state scale. J. Sport Exerc. Psychol. 18, 17–35 (1996).
26. Guo, Z., Xiao, L., Van Toorn, C., Lai, Y., Seo, C.: Promoting online learners' continuance intention. Inf. Manag. 53, 279–295 (2016).
27. Guo, Y.M., Poole, M.S.: Antecedents of flow in online shopping: A test of alternative models. Inf. Syst. J. 19, 369–390 (2009).
28. Kim, H., Suh, K.-S., Lee, U.-K.: Effects of collaborative online shopping on shopping experience through social and relational perspectives. Inf. Manag. 50, 169–180 (2013).
29. Wakefield, R.L., Whitten, D.: Mobile computing: A user study on hedonic/utilitarian mobile device usage. Eur. J. Inf. Syst. 15, 292–300 (2006).
30. Nah, F.F.-H., Eschenbrenner, B., DeWester, D.: Enhancing brand equity through flow and telepresence. MIS Q. 35, 731-A19 (2011).
31. Skadberg, Y.X., Kimmel, J.R.: Visitors' flow experience while browsing a web site. Comput. Human Behav. 20, 403–422 (2004).
32. Webster, J., Ho, H.: Audience engagement in multimedia presentations. DATA BASE Adv. Inf. Syst. 28, 63–77 (1997).
33. Trevino, L.K., Webster, J.: Flow in computer-mediated communication. Communic. Res. 19, 539–573 (1992).
34. Hoffman, D.L., Novak, T.P.: Marketing in hypermedia environment foundations: Conceptual foundations. J. Mark. 60, 50–68 (1996).
35. Zhang, K., Chen, C., Zhao, S., Lee, M.: Compulsive smartphone use: The roles of flow, reinforcement motives, and convenience. ICIS Proc. (2014).
36. Guo, Z., Xiao, L., Seo, C., Lai, Y.: Flow experience and continuance intention toward online learning: an integrated framework. ICIS Proc. 1–21 (2012).
37. Nah, F.F.-H., Eschenbrenner, B., DeWester, D.: Enhancing brand equity through flow and telepresence: A comparison of 2D and 3D virtual worlds. ICIS Proc. (2010).
38. Phang, C.W., Kankanhalli, A.: How do perceptions of virtual worlds lead to enhanced learning? An empirical investigation. ICIS Proc. (2009).
39. Kamis, A., Koufaris, M., Stern, T.: Using an attribute-based decision support system for user-customized products online: An experimental investigation. MIS Q. 32, 159–177 (2008).
40. Koufaris, M.: Applying the technology acceptance model and flow theory to online consumer behavior. Inf. Syst. Res. 13, 205–223 (2002).
41. Visinescu, L.L., Sidorova, A., Jones, M.C., Prybutok, V.R.: The influence of website dimensionality on customer experiences, perceptions and behavioral intentions. Inf. Manag. 52, 1–17 (2015).
42. Ben Mimoun, M.S., Garnier, M., Ladwein, R., Benavent, C.: Determinants of e-consumer productivity in product retrieval on a commercial website: An experimental approach. Inf. Manag. 51, 375–390 (2014).
43. Wang, C.-C., Hsu, M.-C.: An exploratory study using inexpensive electroencephalography (EEG) to understand flow experience in computer-based instruction. Inf. Manag. 51, 912–923 (2014).

44. Lowry, P.B., Gaskin, J.E., Twyman, N.W., Hammer, B., Roberts, T.L.: Taking "fun and games" seriously. J. AIS. 14, 617–671 (2013).
45. Goel, L., Johnson, N.A., Junglas, I., Ives, B.: How cues of what can be done in a virtual world influence learning: An affordance perspective. Inf. Manag. 50, 197–206 (2013).
46. Goel, L., Johnson, N., Junglas, I., Ives, B.: Predicting users' return to virtual worlds: A social perspective. Inf. Syst. J. 23, 35–63 (2013).
47. Lee, Y., Chen, A., Ilie, V.: Can online wait be managed? The effect of filler interfaces and presentation modes on perceived waiting time online. MIS Q. 36, 365–394 (2012).
48. Weniger, S., Loebbecke, C.: Researching cognitive absorption in the context of fun-oriented information systems usage: An exploratory study. ECIS Proc. (2011).
49. Goel, L., Johnson, N.A., Junglas, I., Ives, B.: From space to place: Predicting users' intentions to return to virtual worlds. MIS Q. 35, 749-A5 (2011).
50. Deng, L., Turner, D.E., Gehling, R., Prince, B.: User experience, satisfaction, and continual usage intention of IT. Eur. J. Inf. Syst. 19, 60–75 (2010).
51. Jia, R., Hartke, H., Pearson, J.: Can computer playfulness and cognitive absorption lead to problematic technology usage? ICIS Proc. (2007).
52. Shang, R.-A., Chen, Y.-C., Shen, L.: Extrinsic versus intrinsic motivations for consumers to shop on-line. Inf. Manag. 42, 401–413 (2005).
53. Saadé, R., Bahli, B.: The impact of cognitive absorption on perceived usefulness and perceived ease of use in on-line learning. Inf. Manag. 42, 317–327 (2005).
54. Hess, T.J., Fuller, M. a., Mathew, J.: Involvement and decision-making performance with a decision aid. J. Manag. Inf. Syst. 22, 15–54 (2005).
55. Zhang, H., Lu, Y., Gupta, S., Zhao, L.: What motivates customers to participate in social commerce? The impact of technological environments and virtual customer experiences. Inf. Manag. 51, 1017–1030 (2014).
56. Animesh, A., Pinsonneault, A., Yang, S.B., Oh, W.: An odyssey into virtual worlds. MIS Q. 35, 789–810 (2011).
57. Scott, J.E., Walczak, S.: Cognitive engagement with a multimedia ERP training tool. Inf. Manag. 46, 221–232 (2009).
58. Webster, J., Ahuja, J.S.: Enhancing the design of web navigation systems: The influence of user disorientation on engagement and performance. MIS Q. 30, 661–678 (2006).
59. Bilgihan, A., Nusair, K., Okumus, F., Cobanoglu, C.: Applying flow theory to booking experiences. Inf. Manag. 52, 668–678 (2015).
60. Huang, L.-Y., Hsieh, Y.-J., Wu, Y.-C.J.: Gratifications and social network service usage: The mediating role of online experience. Inf. Manag. 51, 774–782 (2014).
61. Sharkey, U., Acton, T., Conboy, K.: Optimal experience in online shopping: The influence of flow. ECIS Proc. (2012).
62. Theotokis, A., Doukidis, G.: When adoption brings addiction: A use-diffusion model for social information systems. ICIS Proc. 13 (2009).
63. Ha, I., Yoon, Y., Choi, M.: Determinants of adoption of mobile games under mobile broadband wireless access environment. Inf. Manag. 44, 276–286 (2007).
64. Solomon, M.R.: Consumer behavior: Buying, having, and being. Prentice Hall, New York (2014).

Principles in the Design of Mobile Medical Apps: Guidance for Those who Care

Kenny R. Lienhard and Christine Legner

University of Lausanne, Faculty of Business and Economics (HEC), Department of Information Systems, Lausanne, Switzerland
{kenny.lienhard,christine.legner}@unil.ch

Abstract. The promises of mobile technology in healthcare have led to a great many mobile apps in public app stores that target patients with specific illnesses. Medical experts have criticized the status quo of mobile medical apps owing to the low level of professional medical involvement in mobile app design, leading to weak clinical performance and a poor integration of these tools into clinical practice. Grounded in an action design research study, we build and evaluate a mobile app for elderly patients with age-related macular degeneration. We formalize our learnings and provide a set of design principles to guide the effective and feasible construction of mobile medical apps. Our study systematically develops design knowledge that helps to bridge the current gap between the rapid advances in mobile technology and the specific needs of the healthcare sector.

Keywords: mobile health, mobile medical app, mobile patient monitoring, healthcare, design science

1 Introduction

Mobile applications (mobile apps) are seen as a potentially transformative technology that provides individual level support to healthcare consumers and new ways for physicians to partner and interact with their patients. While information used for personal healthcare is traditionally captured via self-reporting surveys and doctor consultations, mobile devices with embedded sensors offer opportunities to establish a continued exchange of information between patients and physicians. Such information exchanges are particularly important concerning patients with chronic illnesses. The promises of mobile health and cutting-edge mobile technologies have led to a great many mobile health apps in public app stores. The two largest mobile platforms, Android and iOS, host more than 165,000 mobile apps on medical topics, of which 9% address topics of screening, diagnosis and monitoring a broad spectrum of illnesses [1]. We focus on this category, and refer to them as *mobile medical apps*.

Innovation in mobile technology has outpaced the critical evaluation of the impact of mobile medical apps [2, 3]. Medical experts have criticized the current state of mobile medical apps, because they are predominantly technology-driven and thus fail to meet the requirements of clinical practice [2]. While existing mobile apps over-

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Lienhard, K.R.; Legner, C. (2017): Principles in the Design of Mobile Medical Apps: Guidance for Those who Care, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1066-1080

emphasize technological aspects, they tend to poorly integrate medical expertise [4]. For instance, only 32% of 63 mobile apps that target patients with colorectal diseases were built based on medical professionals' input on development and content [5]. In another study, a mobile app for melanoma risk assessment available on public app stores was compared with an experienced physician's risk assessment for low-risk and high-risk lesions [6]. The researchers found that the mobile app's sensitivity (i.e. its ability to correctly identify those with a disease) was 20% and its specificity (i.e. its ability to correctly identify those without the disease) 92%. The low involvement rate of medical professionals not only increases the risk that ineffective or even potentially harmful tools will be used by patients, it also leads to a poor acceptance rate of mobile medical apps among physicians and therefore a low integration of these tools into daily clinical practice [7]. Based on these studies, we argue that integrating software developers' and medical professionals' expertise is a prerequisite for designing successful mobile medical apps. However, we still lack knowledge about how to best design mobile medical apps that integrate both aspects. We address this gap in the research with the following research question: *What are suitable principles in mobile medical app design?*

Our research follows a design science research approach "that uses artifact design and construction (learning through building) to generate new knowledge and insights into a class of problems" [8]. In a 30-month action design research (ADR) [9] study, we worked in an interdisciplinary team of researchers and practitioners to develop a mobile medical app for elderly patients with low vision due to cases of age-related macular degeneration. Age-related macular degeneration is a form of sight loss caused by damage to the retina. The result is a shadowlike void in the center of a patient's visual field. We evaluated our artifact in two clinical studies with 124 patients and continuously reflected on learnings, which allowed us to generate design knowledge throughout the process and to refine insights when new inputs became available. We have formulated and transformed our learnings into a generic set of design principles that guide the effective and feasible construction of mobile medical apps. Here, we systematically develop design knowledge that helps to bridge the current gap between the rapid advances in mobile technology and the creation of sustainable mobile app solutions in the healthcare sector.

The paper proceeds as follows. In Section 2, we provide a background on the literature of our study, identify the gap in the research, and formalize the research question. We then describe the research method. Next, we present Alleye – a mobile app that targets patients with age-related macular degeneration. In Section 5, we formalize the learnings from our project and present a set of principles to guide mobile medical app design. Finally, we summarize our paper's contributions, describe the research limitations, and provide an outlook for future research.

2 Background

Mobile medical apps leverage modern mobile devices such as smartphones and tablet computers, built-in sensor technology, and related software development kits (SDKs) to screen, diagnose, and monitor a patient's illness. Screening is the routine examination of individuals for indications of illness or of high risk for illness [10]. Diagnosis is the inferred state that an illness is present in a person [10]. Mobile patient monitoring uses "technology to manage, monitor, and treat a patient's illness from a distance" [11] once an illness has been attributed to an individual.

A number of studies have developed mobile app solutions for particular illnesses, including diabetes [12, 13], asthma [14], and depression [15], and have reported lessons learnt. Goyal et al. [13] take a user-centered design approach, ensuring that the features of a mobile app are informed by the needs of patients with type 2 diabetes. The resulting application allows patients to self-monitor their physical activities, diets, and weights, to identify glycemic control patterns in relation to their lifestyles, and to guide them towards remedial decision-making. Årsand et al. [12] illustrate that their mobile app can motivate type 2 diabetes patients to think about how they can improve their health. The authors conclude that their system has the potential to support the collaboration between patients and clinicians. In another study, Oresko et al. [16] integrate a Holter monitor with mobile technology and develop smartphone-based cardiovascular disease detection. Another study presents a remote monitoring system for elderly patients with multiple chronic conditions [17] that allows users to see current medical reports on their smartphones based on sensor data, to perform new measurements, and to communicate with caregivers via the mobile app. Schnall et al.'s study evaluates existing mobile apps for patients living with HIV and concludes that the design of such mobile apps requires a thoughtful, patient-centered, and evidence-based approach [18].

From a medical perspective, recent healthcare research has revealed that a large number of mobile apps available in public app stores are not based on empirical evidence [19]. These shortcomings can have serious consequences. For instance, Wolf et al. [20] measure the performance of four mobile apps that evaluate photographs of skin lesions. When such a picture is evaluated, the mobile app gives the user feedback about the likelihood of malignancy. The sensitivity of the investigated mobile apps ranged from 6.8% to 98.1%. Hamilton and Brady [4] link the weak performance of some existing mobile apps to low professional medical involvement in the design of mobile apps. Based on the analysis of 111 mobile apps that focus on pain management, one study found that the content of mobile apps contain misleading claims and a lack of academic references [7].

Software technology-oriented communities suggest mobile health frameworks that target developers of mobile apps [21, 22]. One study provides an ontology-based context model and a related application framework that focuses on alarm notification in chronic patient care [22]. Broens et al. [21] suggest a framework to facilitate the use of contextual information (i.e. context acquisition, context provisioning, and context reasoning) for user-tailored mobile apps. While mobile health frameworks are valuable to ensure software component re-usability or functional decomposition, they target the project's implementation phase rather than the design and conceptualization of a mobile medical app. Other studies suggest specific architectural approaches for mobile health app usage [23]. For instance, Kumar et al. [24] performed a comprehensive survey on the use of ubiquitous computing for remote cardiac patient monitoring. They discuss the architecture and quality of service characteristics of the underlying platform for mobile cardiac monitoring systems.

We identified two research opportunities in the scientific literature. First, current studies on mobile medical apps mainly report on the design of illness-specific tools without abstracting higher-level concepts or design principles from their specific solutions. This makes it difficult to apply the learnings from one illness-specific study to the design of a mobile medical app in another context. Second, the rapid advances in mobile technology outpace the rigorous and critical evaluation of the impacts of mobile apps. This leads to a situation in which mobile medical apps continue to proliferate, with little evidence of their effectiveness and little support for understanding how best to design these tools [2]. We build on these gaps in the research and seek to answer the following research question: *What are suitable principles in mobile medical app design?*

3 Research Method

Considering our research goal, we opted for the design science research paradigm, which emphasizes a construction-oriented view of information systems (IS), i.e. research centered around designing and building innovative information technology (IT) artifacts to solve the identified business needs [25]. The research we present here derives rigor from the effective use of the medical and the IS knowledge bases. Our research process followed the ADR approach proposed by Sein et al. [9], a research method for "generating prescriptive design knowledge through building and evaluating ensemble IT artifacts in an organizational setting" (see Figure 1). The ADR project we present here was an engaged research collaboration between academics (two medical and two IS researchers) and practitioners (two senior physicians, a graphic designer, and a software developer). Our ADR study seeks to build and evaluate an innovative IT artifact (i.e. a mobile medical app) and uses heuristic theorizing to synthesize information about artifact solution [26]. The ADR project started in November 2013 with the problem formulation and continued with two building, intervention, and evaluation cycles. The project reached its first complete state in May 2016, when we formalized our learnings from constructing the IT artifact into a set of design principles.

Problem formulation: Our research was driven by the practical need to design a mobile medical app that provides a way for patients to participate in the identification of age-related macular degeneration and the monitoring of this illness. Our case is particularly interesting from both a medical and an IS perspective. From a medical perspective, global projections of any age-related macular degeneration cases are 196 million by 2020, rising to 288 million in 2040 [27]. Current treatment regimens in age-related macular degeneration are suboptimal, owing to 1) the late identification of treatable age-related macular degeneration, 2) non-individualized treatment leading to under-treatment in about 30% of patients, and 3) the challenge to identify the best time for re-treatment after successful treatment and/or treatment suspension. From an IS design perspective, the patients are elderly people who are usually not familiar with touch-based mobile devices; also, owing to their limited vision, it becomes particularly challenging to design an easy-to-use mobile app. Our research effort can be classified as IT-dominant [9], since it emphasizes creating an innovative technological design as its outcome. Our review of extant literature about mobile medical apps revealed that existing studies often report on the design of illness-specific tools, which makes it difficult to transfer their findings when constructing an artifact in another context.

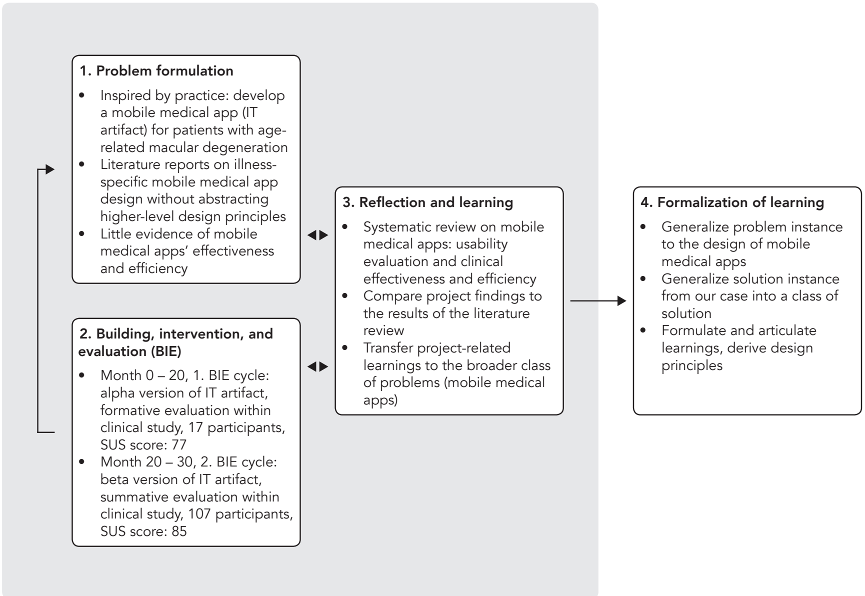


Figure 1. Action Design Research Process (following Sein et al. [9])

Building, intervention, and evaluation (BIE): We built the IT artifact in an agile software development approach with short iterations. During the building phases, we conducted regular meetings (approx. every second week) to evaluate current prototypes in the project team. Each meeting led to new requirements and guided the development of the next prototype version, which we then distributed to members of the ADR team. Overall, we created more than 50 prototypes. Collected data (e.g. field notes from meetings, emails, and visualizations of graphic user interfaces) informed the design of the next prototype version. We also used the prototypes for research purposes, since the data provided us with a history of errors and learnings from the building phases. During the first BIE cycle, we built and evaluated an alpha version of the mobile medical app that included all the functionalities with a strong focus on the measurement task. At that stage, we did not prioritize the presentation and interpretation of the test results in way that is easy to understand by patients. After a project duration of 18 months, we engaged in a naturalistic formative evaluation to determine areas for improvement and refinement of our alpha version. The clinical study took place at the Department of Ophthalmology of the Cantonal Hospital of

Lucerne – Switzerland's largest public eye clinic. This study, in which 17 patients took part, was approved by the Ethics Committee and allowed us to get first-hand feedback from potential end-users. Each patient provided written informed consent. All patients who had had an ophthalmological consultation at the hospital's eye clinic between June and October 2014 were evaluated for inclusion. Patients were excluded if they were unable to use the mobile medical app, i.e. owing to cognitive or visual problems. Patients had a mean age of 78.1 years, and the proportion of women was 53%. Of the participants, 13% used a smartphone daily, while 60% had never used one. During the clinical study, the measurement task had to be executed three times per week, between two routine clinical visits (approx. one month). After data cleaning, we had 240 measurement results.

In the second BIE cycle, we implemented the learnings from the first BIE cycle to build a beta version of the mobile app. The beta version's aim was to be self-explanatory to patients, which required a complete re-design of the user interface and easy-to-understand communication of the measurement results. After several iterations and new prototype versions, we performed a second clinical study as a naturalistic summative evaluation of the IT artifact. We enrolled 107 patients in this study. All patients who had had an ophthalmological consultation at the hospital's eye clinic were evaluated for inclusion. Participants provided written informed consent, and the study was approved by the Ethics Committee. This time, the patients had a mean age of 72.8 years, and the proportion of women was 47%. Of these participants, 40% used a smartphone daily and 53% had never used a smartphone. Compared to the first clinical study, we did not only provide iPod touch devices with the mobile app pre-installed; patients could also use their own smartphone and could download the mobile app via the Apple app store. Of the patients, 29 used their own device. The participants performed repeated self-monitoring measurements between monthly ophthalmological examinations. We collected more than 4,500 measurement results over this clinical study.

In both clinical studies, patients provided oral feedback on the mobile app's user-friendliness, answered pre-defined questions of interest, and filled out the System Usability Scale (SUS). Originally developed by Brooke [28], the SUS is a valid and reliable tool for measuring usability [29]. The SUS has received much attention in the scientific community and, after 30 years from its initial presentation, "has certainly stood the test of time" [29]. In view of the fact that the study participants were elderly patients with substantial visual impairment, and sometimes also impaired cognitive functions, we opted for the short, straightforward SUS as an evaluation instrument. Based on the 500 tools investigated as reference, a SUS score higher than 68 is considered to be above-average [30]. Since the participants in our clinical study were native German speakers, we relied on a German translation of the SUS that was made available via a crowdsourcing project [31].

Reflection and learning: We conducted reflection and learning cycles in parallel to the two BIE cycles. To get an overview of the current state of mobile medical apps, and to learn about their clinical effectiveness and efficiency, we collaborated with medical researchers to perform a comprehensive literature review that integrated the medical and the IS perspectives. We continuously consulted the collected data from the ADR project, i.e. the history of prototypes, presentation material, emails, questionnaire results, opinions from study participants, and measurement results. This allowed us to compare our findings to the literature review results and to apply our situated learning to findings that apply to a broader class of problems. The outcome of the reflection and learning stage was a preliminary set of design principles.

Formalization of learning: Following Sein et al. [9], we distinguish between three levels for this conceptual move. First, we generalized the problem instance to the design of mobile apps that support the monitoring and screening of specific illnesses. Second, we generalized the solutions instance into a class of solution, abstracting highly specific solutions concepts from our own ADR project to make the concepts applicable to the entire class of problems. Third, we captured the knowledge gained in developing an illness-related mobile medical app. Building on the design principles we identified and refined in the previous stage, we fully formulated and articulated our learnings. In the derivation of the design principles, we followed the heuristic theorizing framework suggested by Gregory and Muntermann [26].

4 A Mobile Medical App for Age-related Macular Degeneration

In an engaged academic-practitioner relationship, we built and evaluated Alleye – a mobile app that seeks to provide a way for patients to participate in screening and monitoring age-related macular degeneration. We created this mobile app for Apple's iPhone and the iPod touch, targeting iOS 8.0 and later. It builds on HTML5 technologies and is wrapped inside a web view that provides access to native platform features such as a camera and secure storage. The mobile app has four components: instructions, setup, measurement, and feedback (see Figure 2). For the instructions, Alleye includes a help-center with visual graphics explaining all the functionalities. These graphics were also used by the research assistant to explain the mobile app's use during clinical studies, and patients received a printed booklet with large visual graphics and brief explanations. During initial setup, the patient must insert a unique identification code, so as to match measurement results with his or her electronic health records at the eye clinic. Patients choose to perform the measurement task in training mode or in test mode. The measurement task implemented in Alleye is based on a computerized version of a Vernier hyperacuity alignment task. Hyperacuity is a property of our visual system that allows us to see straight lines as straight. The term derives from the fact that it detects misalignments of borders with a precision that is up to 10 times better than visual acuity. In Alleye, we implemented an alignment task that examines the extent to which the visual system is capable to see straight lines as straight. The performance of an individual completing the measurement task empirically measures the hyperacuity level. Owing to the fact that a drop in hyperacuity precedes a drop in visual acuity, Alleye is capable to detect a decrease of visual acuity before a person sees less. Since this task is monocular (i.e. only the treated eye is open), the patients must select an eye (left or right) before they begin to measure. At the end of each measurement, the patient confirms that the measurement is valid (e.g. that there have been no disturbances). Feedback is provided right after the patient successfully completes a measurement task. The feedback indicates a score and a color scheme inspired by traffic lights, comparing the score to previous test results.

The alpha version of Alleye's user interface followed generic practical guidelines for mobile development. During evaluation in a real-life medical setting, 17 patients reported an average SUS score of 77. From a medical perspective, the first clinical study revealed that the measurement task (i.e. the assessment of the hyperacuity level) is a promising instrument for screening and monitoring patients with age-related macular degeneration. However, we encountered some issues with the mobile app design. For instance, patients could hardly read written text, had difficulties with insufficient color contrast, and the navigation (based on a standard icon-based menu) was unclear to them. To re-design the user interface during the build phase of the second BIE cycle, we simulated the look of the user interface for patients with limited vision and a shadowlike void in the center of their visual field by holding a filter over the mobile device. This provided important insights into the use of colors, minimum font size, or the number of words that should appear on one screen. Further, we implemented a very structured navigation (i.e. minimizing variability via limited navigation options) with buttons occupying the entire screen width in order to guide patients along the mobile app's four components (instruction, setup, measurement, and feedback). The complete redesign of the user interface during the build phase of the second BIE cycle did not impact the measurement task, and captured clinical data could still be compared between the two clinical studies.

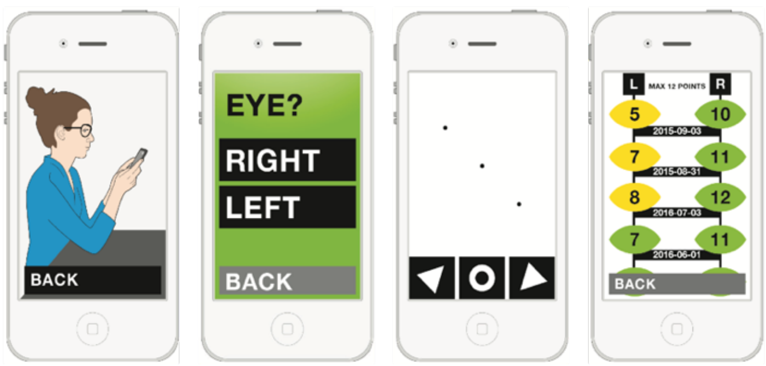


Figure 2. Screenshots of Alleye: Instructions, Setup, Measurement, and Feedback

The second clinical study sought to test the mobile app' use over several months in a real-world medical setting. This would help us to gather data for a longitudinal clinical study and to learn whether or not patients were willing to adopt Alleye and its re-designed user interface. Of 107 patients, 83 provided oral feedback on user-friendliness and filled out the SUS. Compared to the first BIE cycle, the SUS score increased from 77 to 85. We corrected for age and frequent smartphone use in the SUS analysis. In the unadjusted analysis, the estimated mean increase in the SUS score was 8.2 (95% CI: 1.3 to 15.1), while in the adjusted analysis the estimated mean increase in the SUS score only dropped slightly (average SUS score increase 7.5, 95% CI: 0.5 to 14.4, $p = 0.036$). Thus, the user interface optimized for patients with age-related macular degeneration created a higher SUS score. This score was confirmed by oral feedback from patients and a decrease in the help needed by the research assistant to explain the mobile app's usage.

5 Principles in Designing Mobile Medical Apps

Based on the insights gained from the BIE cycles and reviews of other studies on mobile medical apps, we derived four design principles. The design principles' purpose and scope are to provide guidance on how to design mobile medical apps. They capture the knowledge gained about the development of Alleye, and formalize this knowledge to guide the design of other instances of the same class. Justificatory knowledge represents insights from the literature that inform, explain, and validate our design decisions. Table 1 summarizes the design principles, exemplary instantiation based on our research project, and justificatory knowledge.

Table 1. Design Principles and Exemplary Instantiation

Design principles	*Instantiation in Alleye*	*Justificatory knowledge*
DP1: Mobile medical apps should consist of four functional components that guide a patient: instruction, setup, clinical measurement, and analysis and feedback.	• Instruction: help center in mobile app • Setup: identification, select eye for testing • Measurement: alignment task • Analysis and feedback: score and color scheme	[23]
DP2: The user interface should be adapted to cope with patients' physical and cognitive restrictions.	• Limited vision: high color contrast, large font sizes, and buttons that occupy screen width • Cognitive restrictions: limited navigation options, simplify medical information	[32, 33]
DP3: A mobile medical app should build on a robust medical knowledge base, ensuring an evidence-based approach to mobile app design.	• Measurement: hyperacuity level informs diagnosis of age-related macular degeneration • User interface and the use of sensor technology remain independent of the medical knowledge base	[4, 7, 20]
DP4: Mobile medical apps should facilitate both patients' and physicians' routines.	• Instructions trigger conversation between the patient and the physician • Patient feedback should be aligned with actions that the physician can manage and is integrated into his or her routines	[34, 35]

Design principle 1 (DP1): Mobile medical apps should consist of four functional components that guide a patient: instruction, setup, clinical measurement, and analysis and feedback. From our experiences with Alleye, we found that mobile medical apps have four stable functional components that are specific for patient screening and monitoring. Given the logical order of these components, primary navigation elements of mobile medical apps should be structured along these four building blocks in order to provide a patient with guidance. Estrin and Sim [23] present similar components in their seminal paper on an open mobile health architecture, but focus more on technical aspects and leave out the instruction component. The latter is important because it assists the patient through the mobile app's usage and provides explanations on the setup, the clinical measurement task, and the provided feedback. Instructions might be communicated via graphics, audio, and/or videos that support patients who have difficulties reading text. The setup component is crucial to prepare the measurement and to provide context-specific information relevant to perform and analyze the measurement task. For instance, personal data such as weight, age, clinical data, or configuration options might modify the measurement task, are important factors to interpret measurement results, or have diagnostic value (i.e. modify the probability of the presence of the illness). The illness-related measurement is key to the mobile medical app. The analysis and feedback component provides the patient with information about his or her measurement results and might suggest context-specific actions such as the advice to contact a physician. At the same time, feedback should motivate the patient to continue performing clinical measurements. For instance, Marin et al. [36] suggest serious games as a means to keep patients engaged.

Design principle 2 (DP2): The user interface should be adapted to cope with patients' physical and cognitive restrictions. In the development of Alleye, we experienced the limitations of general human interface guidelines provided by mobile platform providers to design mobile apps that target their operating systems. These guidelines provide various user interface patterns, which are helpful to design mobile apps that target a broad audience. However, as we have seen with Alleye, the same human interface guidelines do not consider the very specific limitations of patient groups such as impaired vision, cognitive impairment, or limited motor functions. Designers of mobile medical app should bear in mind end-users' physical and cognitive restrictions [32, 33]. For instance, patients with Parkinson's disease might have difficulties entering data via a smartphone's tiny keyboard. Thus, Parkinson's patients could be provided with structured forms and large buttons to enter data. On the other hand, patients with poor vision can be provided with audio guides and speech recognition instead of written guidance and text fields to enter data. For building the user interface, it is helpful to simulate end-user limitations, so that the designer feels how the mobile app's form and functions works in the hands of future users. In our project, the physicians guided the implementation of this simulation to ensure that designers work with scenarios that are realistic to real-life occurrences. In the case of Alleye, studies on user interface design for elderly people informed the mobile app's development. However, usability studies have certain limitations, since they focus on a user's physical limitations rather than on cognitive restrictions. With

Alleye, we have seen that presenting correct, unbiased information that is hardly understandable by patients not only renders this information useless, but can also cause misunderstandings in patient-physician communication. An evaluation with potential users was required to ensure that patients have the cognitive capabilities to understand the information communicated via the mobile medical app.

Design principle 3 (DP3): A mobile medical app should build on a robust medical knowledge base, ensuring an evidence-based approach to mobile app design. A robust medical knowledge base builds trust among physicians [4, 7], laying the ground for an implementation of mobile medical apps in clinical practice. In Alleye, the assessment of the hyperacuity level informs the diagnosis of age-related macular degeneration. Our mobile app builds on this principle, which is robust. In the course of the use of Alleye, the medical knowledge base might inform us that the changes in hyperacuity are not the same in the various retinal conditions. It might also inform us that, besides hyperacuity, other easily accessible parameters can be measured. While the availability of new sensors embedded in mobile devices might facilitate or even enable the measurement of additional signs and symptoms, they do not necessarily impact the biological model. Advances in mobile technology that impact the mobile app's underlying medical knowledge base would require the mobile app to be clinically reassessed, ensuring that the approach to mobile app design remains evidence-based. While it is very likely that mobile medical apps' forms and functions require adaption and optimization over time owing to technological advances, the underlying biological models and particularly the manifestations (signs and symptoms) of an underlying illness remain fairly stable over time. This is crucial from a medical perspective. Only the stable measurement of clinical parameters allows medical researchers to perform longitudinal clinical studies, and physicians in the hospital can compare a patient's test results over a certain timeframe.

Design principle 4 (DP4): Mobile medical apps should facilitate both patients' and physicians' routines. Patients potentially have a long-term relationship with their physicians. In our project, instructions within Alleye were designed with a specific purpose: They should allow a physician to explain to the patient the mobile app's usage within a few minutes. After basic instructions, the mobile app's use should be self-explanatory. With Alleye, we also learnt that feedback provided after the measurement tasks should be aligned with actions that can be handled by a medical practice. For instance, if the mobile app asks a patient to contact his or her physician because his or her measurement task scores are decreasing, the physician in the medical practice must be aware of the meaning of this call to action. Mobile medical apps offer unique opportunities to improve the quality of the patient-physician relationship [34], since they allow for a continued exchange of clinically useful information that might have remained unrevealed during a routine consultation; such exchanges are of great importance for chronic illnesses in particular. Thus, the design of a mobile medical app for patients is not something that can be done in isolation. Any design of patient routines is linked to the design of the physician's clinical routines. To be fully implemented in routine patient care, it is crucial that a mobile medical app considers requirements from clinical practice. This impacts especially two functional components: instruction (trigger of a patient's routine) and feedback (call to action at the end of a patient's routine). On the one hand, physicians need to explain to the patient the mobile app's use (instructions). This new task should be implemented in existing clinical routines. On the other hand, data collected via a mobile app should become part of the physician's decision-making process. Thus, the mobile app might ask a patient to contact their physician for consultation if their measurement results worsen. The integration of mobile medical apps into clinical routines faces several hurdles, including a lack of knowledge or training on mobile medical apps or incompatibility with current healthcare practices and technology platforms [34]. It is only when physicians adapt their clinical routines to a patient's use of mobile medical apps that the technology's potential can be fully explored.

While DP1 identifies the functional components that guide a patient in using a mobile medical app, DP2 to DP4 provide specific insights on a mobile medical app's architectural design. The identified design principles foster a patient-physician relationship, ensuring that both the patient's and the physician's requirements are addressed when designing a mobile medical app.

6 Conclusion

Mobile medical apps continue to proliferate, with little evidence of their clinical effectiveness and efficiency. From a bottom-up perspective, a number of studies have focused on illness-specific cases of mobile medical apps (e.g. [13–15]). While these studies have revealed important illness-specific insights on building mobile apps in their specific domain, their findings were hardly generalizable to other illnesses. From a top-down perspective, more software technology-oriented approaches provided abstract mobile health frameworks (e.g. [21, 22]) that can be implemented in a broad range of mobile medical apps. This paper links these two research streams by generalizing the solution instance (i.e. a mobile app that targets patients with age-related macular degeneration) into a class of solution. Thus, we abstracted highly specific solution concepts from our project to make the concepts applicable to the entire class of problems (i.e. a mobile medical app that targets patients with a specific illness). Therefore, our study is among the very first to provide principled design knowledge for mobile medical apps. The suggested design principles form a theoretical contribution [26], since they extend the body of knowledge on the creation of mobile app solutions in the healthcare sector. The design principles also assist practitioners in solving current and anticipated problems in the design of mobile medical apps. The abstraction of our learnings allows practitioners to build on the suggested design principles and to apply them in the development of a mobile medical app that targets a different illness than age-related macular degeneration.

Further, our project revealed the importance of involving people from multiple disciplines in a mobile medical app project. In line with our argument, Nilsen et al. [37] criticize current mobile health tools that arise from siloed fields with little reference to previous research. Doing research in an interdisciplinary team involves additional effort, since the team must create and share a vocabulary. However, such an interdisciplinary collaboration enables solutions that could hardly emerge within a single discipline. The application of each of the four suggested design principles in future studies calls for an interdisciplinary collaboration, since they all require inputs from both the medical and the software technology's side.

While our research is grounded in a successful 30-month research project, it has limitations. Although we integrated our insights with findings from the literature to inform, explain, and validate our design decisions, we cannot guarantee that our findings are exhaustive or fully independent of our specific research project. The research we presented here has opened up possibilities for new and exciting future research. Our design principles can serve as a basis to develop a design theory, as suggested by Gregor and Jones [38]. While a design theory provides prescriptions for the design of an artifact, future research should also study de facto implementation and how mobile medical apps change interactions between patients and physicians. What affordances and constraints do mobile medical apps bring to daily clinical practice? And how does the physician's corresponding clinical information system interact with a patient's mobile medical app? These are important issues to address the current gap between mobile technology advances and critical evaluation of the impacts of mobile medical apps in healthcare.

References

1. IMS Institute for Healthcare Informatics: Patient Adoption of mHealth. (2015).
2. Kumar, S., Nilsen, W.J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., Riley, W.T., Shar, A., Spring, B., Spruijt-Metz, D.: Mobile Health Technology Evaluation: The mHealth Evidence Workshop. American Journal of Preventive Medicine. 45, 228–236 (2013).
3. Boulos, M.N.K., Brewer, A.C., Karimkhani, C., Buller, D.B., Dellavalle, R.P.: Mobile Medical and Health Apps: State of the Art, Concerns, Regulatory Control and Certification. Online Journal of Public Health Informatics. 5, (2014).
4. Hamilton, A.D., Brady, R.R.W.: Medical Professional Involvement in Smartphone "Apps" in Dermatology. British Journal of Dermatology. 167, 220–221 (2012).
5. O'Neill, S., Brady, R.: Colorectal Smartphone Apps: Opportunities and Risks. Colorectal Disease. 14, e530–e534 (2012).
6. Robson, Y., Blackford, S., Roberts, D.: Caution in Melanoma Risk Analysis with Smartphone Application Technology. British Journal of Dermatology. 167, 703–704 (2012).
7. Rosser, B.A., Eccleston, C.: Smartphone Applications for Pain Management. Journal of Telemedicine and Telecare. 17, 308–312 (2011).
8. Kuechler, W., Vaishnavi, V.: A Framework for Theory Development in Design Science Research: Multiple Perspectives. Journal of the Association for Information Systems. 13, 395 (2012).
9. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R.: Action Design Research. MIS Quarterly. 35, 37–56 (2011).
10. Miettinen, O.S.: Theoretical Epidemiology: Principles of Occurrence Research in Medicine. Wiley New York (1985).
11. Kay, M., Santos, J., Takane, M.: mHealth: New Horizons for Health Through Mobile Technologies. World Health Organization. 66–71 (2011).

12. Årsand, E., Tatara, N., Hartvigsen, G.: Mobile Phone-based Self-management Tools for Type 2 Diabetes: The Few Touch Application. Journal of Diabetes Science and Technology. 4, 328–336 (2010).
13. Goyal, S., Morita, P., Lewis, G.F., Yu, C., Seto, E., Cafazzo, J.A.: The Systematic Design of a Behavioural Mobile Health Application for the Self-management of Type 2 Diabetes. Canadian Journal of Diabetes. (2015).
14. Burnay, E., Cruz-Correia, R., Jacinto, T., Sousa, A.S., Fonseca, J.: Challenges of a Mobile Application for Asthma and Allergic Rhinitis Patient Enablement – Interface and Synchronization. Telemedicine and e-Health. 19, 13–18 (2013).
15. Burns, M.N., Begale, M., Duffecy, J., Gergle, D., Karr, C.J., Giangrande, E., Mohr, D.C.: Harnessing Context Sensing to Develop a Mobile Intervention for Depression. Journal of Medical Internet Research. 13, e55 (2011).
16. Oresko, J.J., Jin, Z., Cheng, J., Huang, S., Sun, Y., Duschl, H., Cheng, A.C.: A Wearable Smartphone-based Platform for Real-time Cardiovascular Disease Detection via Electrocardiogram Processing. Information Technology in Biomedicine, IEEE Transactions on. 14, 734–740 (2010).
17. Boulos, M.N., Wheeler, S., Tavares, C., Jones, R.: How Smartphones Are Changing the Face of Mobile and Participatory Healthcare: An Overview, with Example from eCAALYX. Biomedical Engineering Online. 10, 24 (2011).
18. Schnall, R., Mosley, J.P., Iribarren, S.J., Bakken, S., Carballo-Diéguez, A., Brown III, W.: Comparison of a User-centered Design, Self-management App to Existing mHealth Apps for Persons Living with HIV. JMIR mHealth and uHealth. 3, (2015).
19. Majeed-Ariss, R., Baildam, E., Campbell, M., Chieng, A., Fallon, D., Hall, A., McDonagh, J.E., Stones, S.R., Thomson, W., Swallow, V.: Apps and Adolescents: A Systematic Review of Adolescents' Use of Mobile Phone and Tablet Apps that Support Personal Management of their Chronic or Long-term Physical Conditions. Journal of Medical Internet Research. 17, (2015).
20. Wolf, J.A., Moreau, J.F., Akilov, O., Patton, T., English, J.C., Ho, J., Ferris, L.K.: Diagnostic Inaccuracy of Smartphone Applications for Melanoma Detection. JAMA Dermatology. 149, 422–426 (2013).
21. Broens, T., Van Halteren, A., Van Sinderen, M., Wac, K.: Towards an Application Framework for Context-aware mHealth Applications. International Journal of Internet Protocol Technology. 2, 109–116 (2007).
22. Paganelli, F., Giuli, D.: An Ontology-based Context Model for Home Health Monitoring and Alerting in Chronic Patient Care Networks. In: Proceedings of the Advanced Information Networking and Applications Workshops, 2007, AINAW'07. 21st International Conference on. pp. 838–845. IEEE (2007).
23. Estrin, D., Sim, I.: Open mHealth Architecture: An Engine for Health Care Innovation. Science. 330, 759–760 (2010).
24. Kumar, S., Kambhatla, K., Hu, F., Lifson, M., Xiao, Y.: Ubiquitous Computing for Remote Cardiac Patient Monitoring: A Survey. International Journal of Telemedicine and Applications. 2008, 3 (2008).
25. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly. 28, 75–105 (2004).
26. Gregory, R.W., Muntermann, J.: Heuristic Theorizing: Proactively Generating Design Theories. Information Systems Research. 25, 639–653 (2014).

27. Wong, W.L., Su, X., Li, X., Cheung, C.M.G., Klein, R., Cheng, C.-Y., Wong, T.Y.: Global Prevalence of Age-related Macular Degeneration and Disease Burden Projection for 2020 and 2040: A Systematic Review and Meta-analysis. The Lancet Global Health. 2, e106–e116 (2014).
28. Brooke, J.: SUS – A Quick and Dirty Usability Scale. Usability Evaluation in Industry. 189, 4–7 (1996).
29. Lewis, J.R., Sauro, J.: The Factor Structure of the System Usability Scale. In: Human Centered Design. pp. 94–103. Springer (2009).
30. Sauro, J.: Measuring Usability with the System Usability Scale (SUS), http://www.measuringu.com/sus.php (Accessed: 17.08.2016).
31. Rummel, B.: Crowdsourcing the Translation of SUS, https://experience.sap.com/skillup/system-usability-scale-jetzt-auch-auf-deutsch/ (Accessed: 29.10.2016).
32. Holzinger, A., Searle, G., Nischelwitzer, A.: On some Aspects of Improving Mobile Applications for the Elderly. In: Universal Access in Human Computer Interaction. Coping with Diversity. pp. 923–932. Springer (2007).
33. Demiris, G., Afrin, L.B., Speedie, S., Courtney, K.L., Sondhi, M., Vimarlund, V., Lovis, C., Goossen, W., Lynch, C.: Patient-centered Applications: Use of Information Technology to Promote Disease Management and Wellness. A White Paper by the AMIA Knowledge in Motion Working Group. Journal of the American Medical Informatics Association: JAMIA. 15, 8–13 (2008).
34. Silow-Carroll, S., Smith, B.: Clinical Management Apps: Creating Partnerships between Providers and Patients. Issue Brief (The Commonwealth Fund). 30, 1–10 (2013).
35. Free, C., Phillips, G., Galli, L., Watson, L., Felix, L., Edwards, P., Patel, V., Haines, A.: The Effectiveness of Mobile-health Technology-based Health Behaviour Change or Disease Management Interventions for Health Care Consumers: A Systematic Review. PLOS Medicine. 10, e1001362 (2013).
36. Marin, J.G., Navarro, K.F., Lawrence, E.: Serious Games to Improve the Physical Health of the Elderly: A Categorization Scheme. In: International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services. Barcelona, Spain (2011).
37. Nilsen, W., Kumar, S., Shar, A., Varoquiers, C., Wiley, T., Riley, W.T., Pavel, M., Atienza, A.A.: Advancing the Science of mHealth. Journal of Health Communication. 17, 5–10 (2012).
38. Gregor, S., Jones, D.: The Anatomy of a Design Theory. Journal of the Association for Information Systems. 8, 312–335 (2007).

Consumer Preferences for Product Information and Price Comparison Apps

Jens Fölting[1], Stephan Daurer[2], and Martin Spann[1]

[1] LMU Munich, Institute of Electronic Commerce and Digital Markets, Munich, Germany
jens.foelting@googlemail.com, spann@spann.de
[2] DHBW Ravensburg, Wirtschaftsinformatik, Ravensburg, Germany
daurer@daurer.net

Abstract. Product information and price comparison apps on smartphones play an increasing role in consumers' purchase decision process. Consumers are able to choose from a variety of product information search applications (apps) which mainly differ with respect to the information that is provided to consumers during their search process. The goal of this study is to analyze preferences regarding different information types that such apps provide. We conduct an adaptive choice-based conjoint analysis combined with a between subject experiment for a sample of 330 consumers. We identify differences between high- and low-involvement products. Individual differences are explained using psychometric latent constructs. Our results reveal heterogeneous preferences which also depend on the product category. Consumer attitudes like quality vs. price consciousness and green consumer values influence the valuation of certain information types.

Keywords: Consumer preferences, adaptive choice-based conjoint analysis, mobile commerce, smartphone applications

1 Introduction

The diffusion of smartphones increasingly improves consumers' means to search and access information online using the mobile internet. For instance, in the domain of product information and price comparison consumers are able to choose among many different smartphone applications (apps) for product information search. Such apps are widely used [1, 2]. We consider product information search apps as smartphone applications that provide information on physical products using the mobile internet. Most of these apps provide barcode scanning, location-based services and different types of product information. Today's product information search apps provide a magnitude of different features that go well beyond basic price comparison [2].
There is some research on app success in general. For instance, Lee and Raghu [3] conduct a survival analysis on 300 apps in Apple's App Store and Yang investigates the acceptance of mobile applications among young Americans [4]. However, to the best of our knowledge, there is no literature on consumer preferences regarding app features that is category specific. In particular, so far there is little known about product information search apps and consumers. The goal of this study is to analyze consumer

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Fölting, J.; Daurer S.; Spann, M. (2017): Consumer Preferences for Product Information and Price Comparison Apps, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1081-1095

preferences regarding different features of product information search apps with a clear focus on different types of product information that an app might provide. We conduct an adaptive choice-based conjoint analysis combined with a between subject experiment for an online sample of 330 consumers. Here we identify differences between the searches for high- versus low-involvement products. Because consumers' preferences are estimated on an individual level, differences can be explained using psychometric latent constructs in a regression analysis.

Towards this goal we work with three research questions that are related to research gaps that we identify: (1) Which types of product information do consumers prefer in their search process using smartphone apps? (2) How does the distribution of the relative attribute importance differ between information searches for high- versus low-involvement products? and (3) Can information search behavior be explained by a certain set of psychometric constructs? The results are relevant for marketers and smartphone app developers alike. Developers will be able to include consumer preferences in their requirements when designing new apps. Marketers might benefit from our insights in their targeting activities. For instance, retailers could use our findings to provide app suppliers with specific product information for inclusion in their app.

We aim to contribute to the literature in different ways. First, we measure consumers' preferences regarding product information search apps and identify "must-have" and "nice-to-have" features. Here, we focus on features that are related to different types of product information. Second, we provide evidence that these preferences and thus app usage behavior depends on the product category. Third, we explain consumer specific differences based on various consumer attitudes.

The remainder of this paper is organized in the following way: We first review related literature to identify research gaps. Then, we outline our data and method. Next, we present the results of our empirical study with the two subsections adaptive choice-based conjoint analysis and regression analysis. Finally, we discuss implications for practitioners as well as for researchers.

2 Related Literature

In consumer search the identification of sellers and the determination of prices are important, as Stigler [5] points out in his seminal paper. However, the search for information (which is required e.g., for price comparisons) is costly and therefore market inefficiencies may be partly explained by search or transaction costs [5]. The internet facilitates information search and it is widely acknowledged that the internet has a major impact on consumer search behavior [6]. Search on the internet even substitutes search in other channels [7]. The diffusion of smartphones adds a new dimension with the mobile internet. While early stage mobile internet search was mainly browser-based [8], today search is usually driven by apps. Mobile devices (smartphones and tablet computers) are used as both research and purchase devices [9]. Since they are portable and provide location-independent access to the internet they reduce consumers' overall search costs [10].

Smartphones support product information search [1]. While there is some research on optimal product information search [e.g., 14, 15] the term product information is manifold and depends on the context. There are various attributes of products that may be relevant in the decision-making process of consumers. An obvious attribute is price. However, Dickson and Sawyer [11] find that not all consumers engage in price comparisons. Nevertheless, price is an important component of offerings. Product quality, which is reflected by different product characteristics, is also a very important attribute [12]. Furthermore, there are consumers that look also on other aspects, for instance information on environment-friendly production processes [13]. Companies' approach to corporate social responsibility is considered by certain consumers [16]. Furthermore, information on sustainability is relevant for some consumers. Winkler von Mohrenfels and Klapper [17] find that relevant mobile product information may increase brand perception and at least for some product categories (e.g., organic foods) it may increase consumers' willingness-to-pay. Finally, consumers might look for reports of other consumers on their user experience. Such information is easily conveyed via the internet in form of consumer reviews or electronic word-of-mouth [18].

As consumers are subjected to a vast quantity of information their judgment eventually could be blurred. This phenomenon is called information overload. It assumes that consumers exhibit "finite limits to the amount of information they can assimilate and process" [19] and that if these limits are exceeded consumers become confused and make poorer decisions. While it is known what technology (i.e. product information search apps) is able to provide and which different types of product information may be relevant in consumer decision making, it remains open which features of such apps (i.e. which product information) are preferred by their users. This leads us to our first research question: (1) Which types of product information do consumers prefer in their search process using smartphone apps?

Based on previous research the importance of different types of product information seems to vary depending on the situation. There is empirical evidence that information choice behavior of consumers differs by product category [1]. This is consistent with earlier reports that in general consumer behavior varies by product category [20]. During information search, price information is more intensely requested by consumers looking for durables compared to consumables [1]. User reviews are more relevant for consumables. User reviews are also more retrieved in product search concerning utilitarian goods, while the price seems to be more important for hedonic goods. Looking at search goods, information on product characteristics is slightly more demanded as opposed to experience or credence goods [1]. In general, product information needs are correlated with the involvement of consumers [21]. Hence, our second research question is: (2) How does the distribution of the relative attribute importance differ between information searches for high- versus low-involvement products?

Previous literature suggests that consumer preferences are heterogeneous. Based on existing literature a set of consumer attitudes can be identified that potentially plays a role in product information search. The attitude of quality consciousness [22] is an indicator of how thoroughly consumers inform themselves about a product before a purchase. Depending on green consumer values consumers might differently assess information on the sustainability of products or their production processes [23]. Furthermore, price consciousness is an indicator of the effort that consumers take to strike a good deal [24]. With our third research question (3) Can information search behavior be explained by a certain set of psychometric constructs? We will analyze various influencing factors on consumer preferences in this context.

In this study we analyze consumer preferences regarding different features of product information search apps with a clear focus on different types of product information. While there is research on app features like mobile technologies, media integration and social network capabilities (e.g., [2]), or research on category independent app characteristics like price of the app or number of apps by the same developer [3], to the best of our knowledge so far there is no previous research that focuses on features that are related to different types of product information in product information search apps.

3 Data and Method

3.1 Study Design

To answer our three research questions we design a modular study consisting of an online discrete choice experiment with a between subjects design. Regarding the first two research questions we apply a form of conjoint analysis. In this conjoint analysis, we represent a situation in which participants are asked to make choices about product information search apps. We conduct an adaptive choice-based conjoint (ACBC) analysis to estimate consumer preferences.

We gather data through an online survey. For the purpose of a comparative analysis to answer the second research question, we design an online questionnaire with two versions. The two versions differ in regards to the focal product – a high-involvement (HI) product (a laptop computer) and a low-involvement (LI) product (an energy drink). Participants are randomly assigned to one of the questionnaire versions. 330 completed questionnaires could be used for the analysis (156 in the HI- and 174 in the LI-group). The sample consists mainly of students and all the relevant socio-demographic characteristics are approximately equally distributed to ensure for a proper comparison between the two experimental groups (high- and low-involvement product).

In addition to the ACBC analysis, we gather socio-demographic and psychographic data by measuring various latent constructs. These constructs are used to explain the results of the ACBC analysis which relate to the third research question. Here, we apply multiple linear regressions with four relative attribute importance values as dependent variables. The relative attribute importance is estimated through the ACBC using Hierarchical Bayes (HB) on an individual level. As independent variables we employ the constructs mentioned above as they cover individual characteristics regarding consumers' search and purchasing processes. All constructs are measured using a 7-point Likert-type scale. Except for opinion seeking, which contains a reverse-coded item, all constructs are coded in a way that higher scores represent higher levels of the construct.

Consumers' importance of quality and their subsequent effort of gathering information about products before making final purchase decisions are measured by the Quality Consciousness scale [22]. The scale Opinion Leaders and Opinion Seekers [25] is limited to the opinion seeking items in our study because we are only interested in whether consumers are potentially influenced by the attitude of others. In order to measure the degree to which the characteristic trait of considering the environmental impact of one's purchase and consumption behavior influences one's information search, the Green Consumer Values scale [23] is included in the survey. The psychometric measurement of consumers' search effort is based on the search costs scale of Srinivasan and Ratchford [26]. We also include the construct Price Consciousness [24] that evaluates the consumers' willingness to spend an extra effort in order to find low prices for a specified product category.

3.2 Adaptive Choice-Based Conjoint Analysis

Choice of Method and Design. Conjoint analysis (CA) can be characterized as a method for determining consumer preferences for products or services which consist of various distinct attributes [27]. In this paper we apply the adaptive choice-based conjoint (ACBC) analysis which achieves to combine the benefits of adaptive and choice-based conjoint procedures [28]. While classical CA or adaptive conjoint analysis (ACA) ask the respondent to either rank the product concepts or rate them on a scale, choice-based conjoint (CBC) analysis realistically imitates the decision process [29]. ACBC analysis makes use of the benefits of ACA as respondents are able to indicate whether a certain attribute level is completely unacceptable and therefore should be excluded from later questions. In addition, it is possible to indicate "must-have" or "unacceptable" attribute levels which leads to a definite in- or exclusion in the rest of the questionnaire.

To identify suitable products for the two settings (high- and low-involvement product), we transfer generic characteristics of high- and low-involvement purchase decisions – i.e. careful versus superficial information processing, systematic information search versus rather casual information reception, high correlation with personality and lifestyle versus a low correlation, decision for the subjective best product versus decision for an acceptable product and a high influence of reference groups versus a low influence – to particular product categories. Especially computer laptops for the high- and energy drinks for the low-involvement category fulfill the above-mentioned requirements. In course of a pretest we reviewed the selected product categories by applying the Involvement with the Product Category scale by Coulter et al. [30]. The obtained test results suggest a proper choice of product categories. On average energy drinks were rated 2.17 on the involvement scale whereas laptops scored an average rating of 4.5.

Adaptive choice-based conjoint analysis consists of three sections: The build your own (BYO) configuration section, the screening section and the tournament section. In the first section respondents are asked to design their ideal product concept by selecting their preferred attribute levels. As the following concepts will consist of attribute levels

that are relatively concentrated around the respondent's preferred levels, the BYO section is used to reduce the error levels during the rest of the survey.
In the second section of the interview, the screening section, the respondent runs through seven screening tasks, each consisting of four product concepts. Accordingly, each attribute level is depicted at least five times. In the screening section the respondents should indicate the product concepts they would potentially use. The section is intended to realize non-compensatory rules. Therefore, it contains so-called "unacceptable" and "must-have" questions. While the respondents answer the questions in the screening task the software identifies attribute levels, which are systematically being avoided or favored. Correspondingly, after the third screening task the respondent is asked whether one of these recognized attribute levels is completely "unacceptable" or a "must-have". According to the respondents' choices all further product concepts shown will satisfy these specifications.
The tournament section is the last and most important part of the survey. Here, the respondent is asked to make a final decision on a set of product concepts that strictly conform to any cut-off ("unacceptable" / "must-have") rules and that are close to their specified product in the BYO section. Correspondingly, participants can now focus on requirements of secondary importance [28].
All the choice tasks are designed with the software package Sawtooth Software SSI Web, except for the two hold-out tasks that were created manually. These tasks serve as a quality indicator [31] and are constant over all questionnaire versions. According to Johnson and Orme [31] these tasks were created using level overlap. They contain the same concepts in a different order. In contrary to the other choice tasks, hold-out tasks are not utilized for the estimation of the relative importance of the attributes [31].

Choice of Attributes and Levels. In this study, we ask participants to evaluate alternative hypothetical product information search apps based on their attributes. These attributes are the different information features provided by the apps. In order to identify the relevant attributes, a two-stage research process was selected. First, we conduct a market analysis on product information search apps to receive a list of information features that are actually being used in different apps. Second, we execute a literature review to validate the importance and influence of these information features from a theoretical perspective. In course of the market analysis we study six product information search apps, namely Barcoo, Check24, Codecheck, Guenstiger.de, Idealo and Redlaser. The consolidation of the results of the initial market analysis and the literature review yields a total of six attributes, which were chosen for the study at hand. All attributes and their levels are shown in table 1.
The first attribute – price – refers to the fact whether the app provides information on the price of the searched product or not. This attribute relates to the fundamental search literature (e.g., [5]). The second attribute is user reviews and the third is neutral product tests. Both exhibit the same levels as price. Riegner [32] highlights the influence of consumer reviews on purchase decisions in a statistically more descriptive manner, for instance stating that the power of reviews is particularly efficient for pricey electronic products like computers. The fourth attribute is product characteristics. It entails detailed information on product characteristics (e.g., quality information or ingredients

for food). This attribute is divided into three levels. First, product characteristics could be “not available”. Second, it could be available in a raw form or third, it could be available in a condensed or illustrative form. In the latter case the information is presented in a way that helps the consumer to evaluate the information at hand. A typical example for condensed product characteristics is a signal similar to a traffic light. For instance, based on the amount of sugar in a comestible the signal shows red (high amount), amber (average amount) or green (low amount). Since such signals provide an orientation to consumers we assume that they are of higher value in the consumers’ decision process compared to raw product characteristics, especially in times of information overload. The fifth attribute is purchase possibility. We differentiate the following three levels: (1) “no information” on the purchase possibility, (2) “only the name of the retailer” or (3) “exact distance to the retailer” (including the name of that store). It could be shown that consumers might change purchase intentions when they receive information on better offers in the vicinity [33]. The last attribute is information on sustainability. Haws et al. [23] for example conceptualize and demonstrate the influence of an environmentally friendly mind-set on consumer preferences.

4 Empirical Results

4.1 Reliability and Validity

Before presenting and discussing the results of the study, an analysis of the goodness of the available data is required. We assess the reliability of the latent constructs by calculating Cronbach’s Alpha separately for the high-involvement (HI) group and for the low-involvement (LI) group: Quality Consciousness (α_{LI} = 0.77; α_{HI} = 0.79), Opinion Seeking (α_{LI} = 0.95; α_{HI} = 0.97), Price Consciousness (α_{LI} = 0.86; α_{HI} = 0.67), Green Consumer Values (α_{LI} = 0.94; α_{HI} = 0.92) and Cost of Search (α_{LI} = 0.83; α_{HI} = 0.88). The analysis of Cronbach’s Alpha reveals that all latent constructs exhibit a high reliability [34].

Next the reliability, predictive validity and internal validity of the ACBC analysis have to be evaluated. Participant’s choices are potentially influenced by several factors, e.g. negligence and lack of interest. To control for potential biases, two identical hold-out tasks were integrated in the survey. The result of the test-retest statistic shows that 88.4 % of all participants in the HI questionnaire and 82.2 % respectively in LI questionnaire chose the same products in both hold-out tasks. Compared to other studies this test-retest validity is high [35].

To assess the predictive validity of the CBC analysis, we also make use of the hold-out tasks. The predictive validity refers to the ability to predict participants’ choices by using the estimated utility parameters [36]. The corresponding validity measure is the hit-rate. The observed choices were compared to the estimated choices. Here, the hit rate is 76.4 % for the LI and 82.7 % for the HI study. Compared to a random hit rate of 33 % (there are three possible choices) such a hit-rate is considered to be high [37].

The hit-rate which depicts the number of correctly estimated choices can also be used to test the internal validity. Here, we predict subjects' responses to the choice tasks used for estimation. The subsequent hit-rates of 83.4 % for the LI and 85.5 % for the HI study show that a large degree of participants' choices is predicted correctly. The internal validity can thus be considered as high [38]. For the purpose of calculating the predictive and internal validity the none-choice option was not included because in the tournament section of the ACBC the none-alternative was not selectable.

4.2 Estimation and Results

In table 1 we depict the estimated part-worth utilities for the attribute levels in the different groups. Those parameters are normalized (zero-centered) HB estimates. These results indicate face validity. Attribute levels with a higher information content provide higher part-worth utilities. One exception is the attribute purchase possibility in case of HI products. A possible explanation for this finding is consumers' tendency to avoid information overload. As the point of time of searching for information on HI products and the point of time of buying those products might not be identical. Due to the fact that normalized part-worth utilities sum up to zero for each and every attribute, negative part-worth utilities simply denote less desired levels. All signs and therefore the direction of the impact of these utility parameter estimates on the overall utility are plausible.

In contrast to the part-worth utilities of the single attribute levels, a direct interpretation of the superordinate attributes cannot be accomplished [39]. Therefore the measurement of the relative attribute importance is used. The relative attribute importance measures the relevancy of one attribute utility compared to the sum of all attribute utility ranges. To calculate the relative attribute importance, the part-worth utility ranges of each attribute are used. Their part-worth utility ranges are the difference between the highest and the lowest part-worth utility parameter of each attribute. The relative attribute importance of the attribute price is calculated by the utility range of the price divided by the sum of all attribute utility ranges. The same applies to the relative importance of all other attributes. The relative importance for each attribute is depicted in figure 1.

In order to test whether the information search behavior differs significantly for LI and HI products we first compare the variances of each attribute importance. Then we apply a two sample t-test with either equal or unequal variances. While there is no significant deviation of the mean values of the relative attribute importance for the information type price and information on sustainability, all other mean relative attribute importance values differ significantly. Non-parametric Wilcoxon-Mann-Whitney-tests provide congruent results. Thus, these results indeed indicate search behavior to be significantly different for the two product categories.

In both studies price information (34.5 % LI; 33.2 % HI) is of the highest importance for respondents. Likewise product characteristics (24.7 % LI; 27.1 % HI) can be identified to be the second most important feature. The rest of the priority order of the attributes differs in the two groups. In context of information search for LI products participants indicate purchase possibility to be third in the ranking of priority with

15.2 %. This attribute only receives the fifth place in the HI group. This difference might originate from the larger share of the travel costs in the overall price of the product to be purchased.[1] Hence, information on the exact location of a retailer becomes more important the less expensive a product is (ratio of travel costs and price is higher). This interpretation is supported by Kahneman and Tversky's prospect theory. In their behavioral theory they show that consumers perceive decision outcomes as gains and losses, which are defined relative to a certain reference point [40]. As the amount of travel costs increases the overall price (product price plus travel costs) increases, too. In prospect theory travel costs represent losses. In this context a fixed amount of loss x is relatively high for cheaper products.

Table 1. Part-worth utilities (=PWU; HI = High-involvement; LI = Low-involvement product)

Attribute	*Levels*	*PWU HI*	*PWU LI*
Price	Available	99.50	102.90
	Not available	-99.50	-102.90
User reviews	Available	36.06	26.93
	Not available	-36.06	-26.93
Neutral product tests	Available	39.53	28.07
	Not available	-39.53	-28.07
Product characteristics	Condensed information	50.88	49.38
	Raw information	43.22	34.11
	No information	-94.10	-83.50
Purchase possibility	Exact distance to the retailer	9.41	29.14
	Only the name of the retailer	15.43	14.13
	No information	-24.84	-43.27
Information on sustainability	Available	17.61	16.79
	Not available	-17.61	-16.79

A relatively high priority in context of information search is obtained by the attributes user reviews (HI: 12.2 %; LI 9.5 %) and neutral product tests (HI: 13.2 %; LI: 9.5 %). It is quite interesting that the values for the relative importance for these two attributes are that close to each other. This could imply that customers regard user reviews as substitutes for neutral product test reports and vice versa.

In both studies the field of the relative attribute importance is tailed by information on sustainability. Although 59.2 % in the LI study and 69.2 % in the HI study choose to integrate that information feature in the BYO section, the attribute comes up with a relative importance of only 6.2 % in the HI and 6.7 % in the LI group respectively. Correspondingly, consumers value other attributes higher when it comes to a trade-off situation. This finding is confirmed by the results of the screening section. Only 3.6 % of the respondents of the LI and 3.1 % of the respondents of the HI group classify the fact of missing information on sustainability as being unacceptable. Accordingly, for merely 1.6 % (LI group) and 0.7 % (HI group) information on sustainability must be available. Here, it becomes obvious that information on sustainability is rather a "nice-

[1] Please note that here overall price reflects product price plus transaction costs.

to-have" feature. We also identify hygiene factors, e.g., price and product characteristics. Over 96 % of the respondents in both groups choose these two features to be a part of their ideal product information search app. Additionally, the screening section reveals the absence of one of these features to be unacceptable for the majority of the participants.

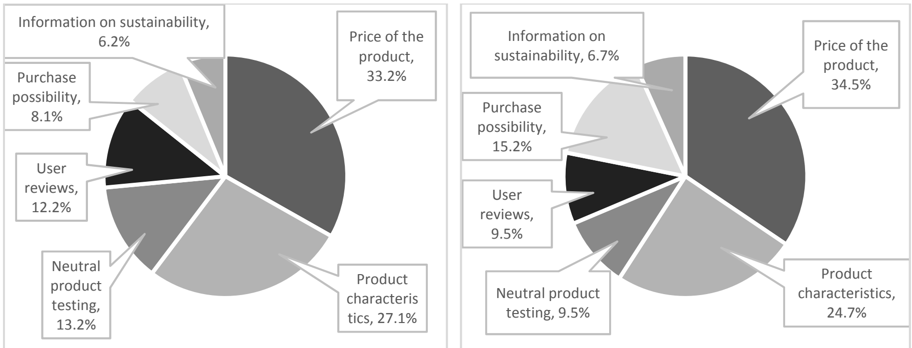


Figure 1. Relative attribute importance (left: HI product; right: LI product)

Subsequently, we analyze the four most important relative attribute importance of the ACBC analysis as a function of the psychometric constructs that we have measured. The relative attribute importance (RAI) provides insights on consumers' attitude towards the different information features while the psychometric constructs measure individual traits concerning consumers' search and purchasing processes. We use multiple linear regressions to explain the relative attribute importance by several psychometric constructs. We estimate the following linear model:

$$\begin{aligned} RAI_{ij} = {} & \beta_{0j} + \beta_{1j} \times quality\ consciousness_i + \beta_{2j} \times \\ & opinion\ seeking_i + \beta_{3j} \times price\ consciousness_i + \beta_{4j} \times \\ & green\ consumer\ values_i + \beta_{5j} \times cost\ of\ search + \beta_{6j} \times \\ & product\ type + \beta_{7j} \times smartphone\ owner_i + \beta_{8j} \times \\ & experience\ with\ product\ information\ search\ app_i + \varepsilon_{ij} \end{aligned} \quad (1)$$

where the dependent variable represents the relative attribute importance of the attributes price, product characteristics, user reviews and neutral product tests. The betas are the estimators of the independent variables of participant *i*. Index *j* reflects the attribute (price, product characteristics, etc.).

First, we test for violations of the assumptions of the linear regression model. To test if our model has a proper functional form, we use the Ramsey test. The results indicate that our models are specified properly. The Breusch-Pagan/Cook-Weisberg test indicates that heteroskedasticity is a problem, however. We therefore use models with robust standard errors [41]. A test for the normality of residuals indicates that the residuals of our models are normally distributed. It is assumed that the residuals are identically and independently distributed. Finally, multicollinearity is not a problem since the maximum variance inflation factor (VIF) is 2.01.

Table 2. Linear regression results

Independent Variable	*RAI price*	*RAI product characteristics*	*RAI user reviews*	*RAI neutral product tests*
Quality Consciousness	-0.015**	0.016***	0.0001	0.005
	(0.006)	(0.005)	(0.004)	(0.003)
Opinion Seeking	-0.003	-0.003	0.003	0.002
	(0.005)	(0.004)	(0.003)	(0.002)
Price Consciousness	0.018***	-0.006	-0.002	-0.003
	(0.005)	(0.004)	(0.002)	(0.002)
Green Consumer Values	-0.013**	0.010**	-0.009***	-0.003
	(0.005)	(0.004)	(0.003)	(0.002)
Cost of Search	0.004	-0.004	0.003	-0.003
	(0.004)	(0.003)	(0.002)	(0.002)
Product Type	-0.021	0.028	0.023**	0.031***
[0 = LI; 1 = HI]	(0.018)	(0.016)	(0.010)	(0.008)
Smartphone Owner	-0.027	-0.002	0.016	0.024***
[0 = no; 1 = yes]	(0.019)	(0.016)	(0.011)	(0.009)
Experience	-0.0004	-0.008	0.002	-0.007
[0 = no; 1 = yes]	(0.014)	(0.012)	(0.008)	(0.006)
Intercept	0.440***	0.166***	0.106***	0.066***
	(0.042)	(0.038)	(0.024)	(0.021)
R^2	0.098	0.081	0.095	0.142
F-test	4.48***	3.89***	4.82***	8.19***
Observations	330	330	330	330

*Robust standard errors in parentheses; *** p<0.01, **p<0.05, * p<0.1*

The F-tests depict the overall significance of the particular regression models. All models are highly significant (see table 2). The dummy product type shows that the category of the inspected product (HI vs. LI) plays a significant role for user reviews (+0.023, $p<0.05$) and neutral product tests (+0.031, $p<0.01$). The special importance of user reviews and neutral product tests in case of HI products can be explained by two facts. First, consumers are interested in reducing their perceived risk by gathering as much information as possible and second, they strive for aligning their purchase decisions with the interest of their reference groups. Both information features user reviews and neutral product tests, help to satisfy this need. When looking at price as an information feature of product information search apps quality consciousness (-0.015, $p<0.05$), price consciousness (+0.018, $p<0.01$) and green consumer values (-0.013, $p<0.05$) are significant. Whereas price consciousness has a positive impact on the relative importance of the attribute price, the constructs quality consciousness and green consumer values exhibit a negative influence. This means that people who are willing to spend an additional effort in order to find low prices for a certain product are more inclined to use the price information during their search process. This seems to be plausible as the primary concern and objective of price conscious consumers is to detect and realize relatively low prices in the process of their buying decisions. Contrary to that, opposite signs of quality consciousness (+0.016, $p<0.01$) and green consumer values (+0.01, $p<0.05$) can be observed in the model for the product characteristics. People that consider the environmental impact of their purchase and consumption behavior are less motivated to use the information feature price. Price is not their primary decision criterion. In fact, highly environmentally conscious people even seem to omit the information feature price to not be distracted from their main intention: to make purchase decisions with integrity. This finding also proposes that the price of a product does not primarily effect their purchase decision. Consequently, they exhibit a higher willingness-to-pay if the product satisfies their needs for sustainability. The same relation is observed for quality-conscious consumers. Furthermore, ecologically-thinking consumers refrain from implying user reviews in their search process (-0.009, $p<0.05$). Ecologically thinking consumers might be less influenced by the opinions of others. Whereas, prior experience with product information search apps is not significant in all models, smartphone ownership has a significant positive influence on neutral product tests (+0.024, $p<0.01$). Despite the wide diffusion of smartphones, the smartphone owner group might still include the more solvent consumers. Neutral product test reports are typically not available for free in contrast to user reviews in form of user-generated content.

5 Conclusion

In this paper we analyze consumers' preferences regarding features of product information search apps with a clear focus on different types of information using an adaptive choice-based conjoint analysis. Price information and product characteristics are the most important features across both product categories. The preferences for other features depend on the product category of the focal product. When searching information on high-involvement products neutral-product test information is important. For low-involvement products consumers value information on purchase possibilities more.

As we estimate consumers' preferences on an individual level, we are able to explain differences using psychometric constructs in a linear regression. Consumer attitudes like quality consciousness, price consciousness and green consumer values influence the valuation of app features. For instance, for quality conscious consumers or consumers with high green consumer values price information is less important. Ecologically-thinking consumers prefer quality information and information on sustainability over user-generated content, such as user-reviews.

Our findings provide various implications for practitioners. With product information search apps consumers are not only increasingly better able to conduct mobile price comparisons they are also better informed about product characteristics. Marketers might provide specific types of product information depending on the product category and depending on consumers' characteristics. App users could be targeted with dedicated mobile advertising [42]. They are not only more likely to find a better alternative offer but they are also more likely to react to the advertising [43]. Furthermore, app developers might include these findings on consumer preferences in their requirements when designing new product information search apps or refining existing ones. As this preference information translates well into real app success in major app stores, those findings are highly relevant for developers as well.

As most research, this study also comes with some limitations: For example, we collect our data using a survey. Therefore our analysis is based on stated preferences which are not consequential for consumers and therefore could lead to a hypothetical bias. However, stated preferences can be used to predict consumer behavior to a significant extent [44]. In addition, the attributes of our ACBC analysis are not exhaustive. The multitude of real apps (across different categories) encompasses many more available features (e.g., barcode scanning vs. manual search). Finally, our survey is based on one country. These limitations provide some avenues for future research. It would be interesting to know if the results are different for other product types (e.g., food vs. electronics). Furthermore, aspects such as customer experience or consumer interaction inside a store when using a product information and price comparison app could be analyzed in further research.

References

1. Daurer, S., Molitor, D., Spann, M., Manchanda, P.: The Impact of Smartphones, Barcode Scanning, and Location-Based Services on Consumers' Search Behavior. In: Proceedings of the 34th International Conference on Information Systems (ICIS) (2013)
2. Engelsma, J., Jumah, F., Montoya, A., Roth, J., Vasudevan, V., Zavitz, G.: Shop Social: The Adventures of a Barcode Scanning Application in the Wild. In: Akan, O., Bellavista, P., Cao, J., Dressler, F., Ferrari, D., Gerla, M., Kobayashi, H., Palazzo, S., Sahni, S., Shen, X. et al. (eds.) Mobile Computing, Applications, and Services, 110, pp. 379–390. Springer, Berlin, Heidelberg (2013)
3. Lee, G., Raghu, T.S.: Determinants of Mobile Apps' Success: Evidence from the App Store Market. Journal of Management Information Systems 31, 133–170 (2014)
4. Yang, H.C.: Bon Appétit for Apps: Young American Consumers' Acceptance of Mobile Applications. Journal of Computer Information Systems 53, 85–96 (2013)
5. Stigler, G.J.: The Economics of Information. Journal of the Political Economy 69, 213–225 (1961)
6. Kuruzovich, J., Viswanathan, S., Agarwal, R., Gosain, S., Weitzman, S.: Marketspace or Marketplace? Online Information Search and Channel Outcomes in Auto Retailing. Information Systems Research 19, 182–201 (2008)
7. Ratchford, B., Lee, M.-S., Talukdar, D.: The Impact of the Internet on Information Search for Automobiles. Journal of Marketing Research 40, 193–209 (2003)
8. Church, K., Smyth, B., Cotter, P., Bradley, K.: Mobile Information Access. A Study of Emerging Search Behavior on the Mobile Internet. ACM Transactions on the Web 1, 1–38 (2007)

9. Ghose, A., Han, S.P., Xu, K.: Mobile Commerce in the New Tablet Economy. In: Proceedings of the 34th International Conference on Information Systems (ICIS), pp. 1–18 (2013)
10. Ghose, A., Goldfarb, A., Han, S.P.: How Is the Mobile Internet Different? Search Costs and Local Activities. Information Systems Research 24, 613–631 (2013)
11. Dickson, P.R., Sawyer, A.G.: The Price Knowledge and Search of Supermarket Shoppers. Journal of Marketing 54, 42–53 (1990)
12. Zeithaml, V.A.: Consumer Perceptions of Price, Quality, and Value. A Means-End Model and Synthesis of Evidence. Journal of Marketing 52, 2–22 (1988)
13. Laroche, M., Bergeron, J., Barbaro-Forleo, G.: Targeting Consumers who are Willing to Pay More for Environmentally Friendly Products. Journal of Consumer Marketing 18, 503–520 (2001)
14. Häubl, G., Dellaert, B., Donkers, B.: Tunnel Vision. Local Behavioral Influences on Consumer Decisions in Product Search. Marketing Science 29, 438–455 (2010)
15. Branco, F., Sun, M., Villas-Boas, J.M.: Optimal Search for Product Information. Management Science 58, 2037–2056 (2012)
16. Peloza, J., Shang, J.: How Can Corporate Social Responsibility Activities Create Value for Stakeholders? A Systematic Review. Journal of the Academy of Marketing Science 39, 117–135 (2011)
17. Winkler von Mohrenfels, H., Klapper, D.: The Influence of Mobile Product Information on Brand Perception and Willingness-to-Pay for Green and Sustainable Products. In: Proceedings of the 33rd International Conference on Information Systems (ICIS), pp. 1–17 (2012)
18. Dellarocas, C.: The Digitization of Word of Mouth. Promise and Challenges of Online Feedback Mechanisms. Management Science 49, 1407–1424 (2003)
19. Malhotra, N.K., Jain, A.K., Lagakos, S.W.: The Information Overload Controversy. An Alternative Viewpoint. Journal of Marketing 46, 27–37 (1982)
20. Bellenger, D.N., Robertson, D.H., Hirschman, E.C.: Impulse Buying Varies by Product. Journal of Advertising Research 18, 15–18 (1978)
21. Schmidt, J.B., Spreng, R.A.: A Proposed Model of External Consumer Information Search. Journal of the Academy of Marketing Science 24, 246–256 (1996)
22. Völckner, F.: The Dual Role of Price: Decomposing Consumers' Reactions to Price. Journal of the Academy of Marketing Science 36, 359–377 (2008)
23. Haws, K.L., Winterich, K.P., Naylor, R.W.: Seeing the world through GREEN-tinted glasses: Green consumption values and responses to environmentally friendly products. Journal of Consumer Psychology 24, 336–354 (2014)
24. Wakefield, K.L., Inman, J.J.: Situational Price Sensitivity: The Role of Consumption Occasion, Social Context and Income. Journal of Retailing 79, 199–212 (2003)
25. Flynn, L.R., Goldsmith, R.E., Eastman, J.K.: Opinion Leaders and Opinion Seekers. Two New Measurement Scales. Journal of the Academy of Marketing Science 24, 137–147 (1996)
26. Srinivasan, N., Ratchford, B.T.: An Empirical Test of a Model of External Search for Automobiles. The Journal of Consumer Research 18, 233–242 (1991)
27. Green, P.E., Srinivasan, V.: Conjoint Analysis in Consumer Research. Issues and Outlook. Journal of Consumer Research 5, 103–123 (1978)

28. Giessmann, A., Stanoevska-Slabeva, K.: What are Developers' Preferences on Platform as a Service? An Empirical Investigation. In: 46th Hawaii International Conference on System Sciences (HICSS), pp. 1035–1044 (2013)
29. Green, P.E., Krieger, A.M., Wind, Y.: Thirty Years of Conjoint Analysis. Reflections and Prospects. Interfaces 31, 56–73 (2001)
30. Coulter, R.A., Price, L.L., Feick, L.F.: Rethinking the Origins of Involvement and Brand Commitment: Insights from Postsocialist Central Europe. Journal of Consumer Research 30, 151–169 (2003)
31. Johnson, R. and Orme, B.K.: Including Holdout Choice Tasks in Conjoint Studies, https://www.sawtoothsoftware.com/download/techpap/inclhold.pdf
32. Riegner, C.: Word of Mouth on the Web: The Impact of Web 2.0 on Consumer Purchase Decisions. Journal of Advertising Research 47, 436 (2007)
33. Daurer, S., Molitor, D., Spann, M.: Measuring Individual Search Costs on the Mobile Internet. In: Proceedings of the 20th European Conference on Information Systems (ECIS) (2012)
34. Peterson, R.A.: A Meta-Analysis of Cronbach's Coefficient Alpha. Journal of Consumer Research 21, 381–391 (1994)
35. Huber, J., Wittink, D., Fiedler, J.A., Miller, R.: The Effectiveness of Alternative Preference Elicitation Procedures in Predicting Choice. Journal of Marketing Research 30, 105–114 (1993)
36. Akaah, I.P., Korgaonkar, P.K.: An Empirical Comparison of the Predictive Validity of Self-Explicated, Huber-Hybrid, Traditional Conjoint, and Hybrid Conjoint Models. Journal of Marketing Research 20, 187–197 (1983)
37. Lenk, P.J., DeSarbo, W.S., Green, P.E., Young, M.R.: Hierarchical Bayes Conjoint Analysis. Recovery of Partworth Heterogeneity from Reduced Experimental Designs. Marketing Science 15, 173–191 (1996)
38. Moore, W.L.: A Cross-validity Comparison of Rating-based and Choice-based Conjoint Analysis Models. International Journal of Research in Marketing 21, 299–312 (2004)
39. Train, K.: Discrete Choice Methods with Simulation. Cambridge University Press, New York (2003)
40. Kahneman, D., Tversky, A.: Prospect Theory. An analysis of decision under risk. Econometrica 47, 263–291 (1979)
41. White, H.: A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica 48, 817–838 (1980)
42. Luo, X., Andrews, M., Fang, Z., Phang, C.W.: Mobile Targeting. Management Science 60, 1738–1756 (2014)
43. Goh, K.Y., Chu, H., Soh, W.: Mobile Advertising. An Empirical Study of Advertising Response and Search Behavior. In: Proceedings of the 30th International Conference of Information Systems (ICIS), vol. Paper 150vol. , pp. 1–17 (2009)
44. Louviere, J.J., Hensher, D.A., Swait, J.D.: Stated Choice Methods. Analysis and Applications. Cambridge Univ. Press, Cambridge (2010)

How to Gauge the Relevance of Codes in Qualitative Data Analysis? – A Technique Based on Information Retrieval

Alexander Keller[1]

[1] University of Passau, Chair of Information Systems (Information- and IT-Service-Management), Passau, Germany
alexander.keller@uni-passau.de

Abstract. Qualitative research has experienced broad acceptance in the IS discipline. Despite the merits for exploring new phenomena, qualitative methods are criticized for their subjectivity when it comes to interpretation. Therefore, research mostly emphasized the development of criteria and guidelines for good practice. I present an approach to counteract the issue of credibility and traceability in qualitative data analysis and expand the repertoire of approaches used in IS research. I draw on an existing approach from the information science discipline and adapt it to analyze coded qualitative data. The developed approach is designed to answer questions about the specific relevance of codes and aims to support the researcher in detecting hidden information in the coded material. For this reason, the paper contributes to the IS methodology with bringing new insights to current methods by enhancing them with an approach from another discipline.

Keywords: qualitative data analysis; information retrieval; relevance of codes; qualitative methods; information science.

1 Introduction

Qualitative research is an accepted methodology in the information system discipline and since a significant growth in 2005, one can see a stable amount of publications in information system (IS) journals like Management Information Systems Quarterly (MISQ), Journal of Management Information Systems (JMIS) and Journal of the Association for Information Systems (JAIS) [1]. Besides studies which use qualitative research methods to answer specific research questions (e.g. [2–4]), much has been written about recommendations and guidelines regarding how qualitative research is conducted within the IS discipline (e.g. [5–7]). Especially in the IS field qualitative research and the findings it generates often have been seen as a minor discipline accompanied with biases regarding the common quality standards [8]. Therefore qualitative research has a strong need to declare itself and reply critics with its own quality criteria (e.g. [9]).

In their MISQ guest editorial on qualitative research in IS Sarker et al. (2013) introduce the principle of transparency and state that "there is a need for clarity in the logic underlying data analysis" [10]. This paper follows their advice and adds a standardized

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Keller, A. (2017): How to Gauge the Relevance of Codes in Qualitative Data Analysis? - A Technique Based on Information Retrieval, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1096-1110

technique to the repertoire of approaches that are used to analyze data in qualitative IS research. Therefore, I have applied an approach from the information science domain and adopted it for the use in qualitative research. With the technique results are generated that grant additional and hidden insights regarding the relevance of codes derived from textual data. This new technique is called *code appearance and relative frequency index* (CARFI). Besides this methodical contribution the technique especially contributes to the field of IS, where relevance aspects are of special interest when studying technical and behavioral issues regarding the role of information technology (IT) (see actual studies like [11, 12]).

The underlying research process started with the need for an appropriate approach to investigate the relevance of codes and is structured in three phases. In the first orientation phase, the qualitative research discipline as well as the IS discipline have been screened for techniques. Due to no appropriate result the scope has been widened to other disciplines. With finding the information retrieval approach the third phase started that aims to adopt the technique for the usage with qualitative data. The results of this phase are presented in this work.

The paper is structured as follows: First, it is explained which methods, data collection techniques and modes of analysis are used in qualitative research and how they collude in the IS discipline. Besides that, I am looking at the coding process and the different techniques used. This is supplemented with existing quality criteria and shortcomings in regard to the interpretation of qualitative data that motivates the development of new analysis procedures. Based on this, I propose a new technique that helps to identify the relevance of codes in textual data sources. I illustrate the technique with real data from interviews with experts. The paper concludes in a discussion of the presented approach and gives an outlook towards future research.

2 Qualitative Research in IS

In the IS domain qualitative research has a long history and aims to study managerial and organizational issues related to innovations in information and communication technologies [13]. Qualitative research methods have its origin in social science and were developed to understand cultural and sociological phenomena [13]. In contrast to quantitative research the focus of qualitative techniques lies on the observation of complex phenomena and situations which cannot be grasped with quantitative measures. Besides qualitative research being used as the only instrument for studying a research topic, it can be combined with quantitative techniques in the same inquiry in a mixed methods approach. Because of both methodologies having different strengths they can rest upon each other to generate richer results in combination compared to being used alone [14, 15].

2.1 Methods, Data Collection and Modes of Analysis

In qualitative research one distinguishes between various methods which specify the research design and strategy. In Table 1 we shortly present five approaches which are widely used in the IS domain and give references which explain the methods from a theoretical perspective (for the information given in Table 1 also see [13] and [16]).

Table 1. Qualitative Research Methods

Method	Description	References
Action Research	The method is a design-orientated approach and aims to widen the stock of knowledge in the specific domain. The goal is the development of solutions in a collaborative manner. Therefore, several cycles consisting of analysis, action and evaluation steps are needed.	[17–19]
Case Study Research	With case studies, complex phenomena are investigated within their real-life context. The approach focuses on cases which represent instances of a particular phenomenon. Therefore, each regarded case can be seen as a research object, which explicates and illustrates the research subject.	[20–23]
Ethnography	The behavioral method aims to generate findings through participatory observation within the direct environment of the research subject. Therefore, a strong and intensive involvement of the researcher is required to study the phenomena in its social context.	[24–26]
Grounded Theory	Grounded theory is a behavioral method with the goal to systematically develop theory from the interaction between data gathering and analysis. Therefore, the approach consists of a process with several rounds of coding (open, axial, selective) to analyze data in which the theory is grounded.	[27, 28]
Qualitative Content Analysis	Qualitative content analysis is based on an iterative process, which is designed for coding data in relation to a category system. The codes, also called categories, of this system are developed in an inductive (from the information sources) and/or deductive (theory orientated) manner.	[29, 30]

Besides the presented methods of qualitative research, different techniques of data collection are distinguished to gather empirical material for the purpose of data analysis. Qualitative material covers mostly non-numeric information which is collected directly in form of interviews, observations and fieldwork or indirectly from secondary sources like published and unpublished documents [13]. This results in textual information sources like interview transcripts and case descriptions which serve as the foundation for further analysis.

The non-numeric nature of qualitative data results in different modes of analysis compared to quantitative methods. In qualitative research it is essential to understand and interpret the given information in regard to the underlying research questions to ensure that conclusions are drawn in the right context. Therefore, the textual data is processed and text passages are coded with a system of categories to create a foundation for analysis and interpretation [31]. Besides approaches which result from the methodology perspective (e. g. grounded theory), there exist additional modes of analysis that can also be seen as philosophies for interpreting qualitative information. Among others the hermeneutic idea is a way of understanding and interpreting textual information and aims to create an understanding for the "[…] underlying coherence or sense" [32]. Therefore, both understanding the whole text on the one side and interpreting the different parts of it on the other side is relevant for analyzing the given information [33]. The approach targets the understanding of the research subject in general and focusses on the relation between different objects on the basis of textual information.

The semiotic approach aims to investigate the implicit and explicit meaning of texts and is divided into three different categories - the content, conversations and discourse analysis [13]. With content analysis the researcher structures the textual information sources and identifies patterns in regard to the given context [34]. In conversation analysis the background of social interaction and behavior is uncovered with the researcher immersing in the particular situation [25]. Discourse analysis "[…] concerns itself with critically analyzing language […] in the context of social interactions" [35]. With this concept not only the provided content of a person, but also their way of talking is investigated.

Figure 1 shows a summary of methods, data collection techniques and modes of analysis used in IS based on Myers (1997)[13]. As the three different parts are coherent

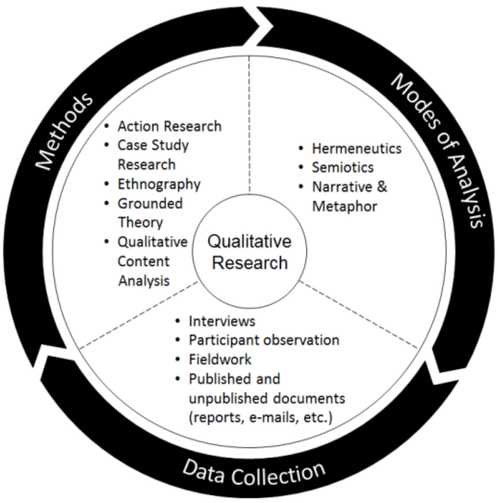


Figure 1. Qualitative Research

and mutually dependent on one another, qualitative research designs evolve from a combination of these aspects.

2.2 Codes and Coding Techniques for Qualitative Data Analysis

With the occurrence of grounded theory methodology the technique of coding emerged in the late 60s and is now said to be one of the most popular techniques of data analysis [36]. Referring to Sarker et al. (2012) about 60% of qualitative studies in the IS discipline use coding techniques to analyze empirical data [1]. Coding is a procedure to structure text and aims to mark segments of textual raw data that contain a specific information in regard to the underlying research question [37]. In the coding process the researcher identifies relevant information in the textual raw data (e.g. interview transcripts, reports, etc.) and attaches an existing or new code to a text segment [38]. A code represents the information given in the data and functions as an empirical generalization of relevant information in regard to the research question.

Looking at code-based qualitative data analysis Gläser and Laudel (2013) distinguishes between two important variations. First the difference in coding themes or coding content and second "[…] the extent to which preexisting theory is used in the coding process […]" [36].
Typically coding results in indexing themes, i.e. a set of codes that "[…] represents the structure of raw data in the text […]" [36]. The intention is to capture the information given in a text segment. In contrast coding content focusses on the content that should be expressed with a particular statement. In this case a code is used as a representation for a typical phenomenon of interest [36]. Besides that, codes can directly be generated from the textual raw data without considering existing theoretical aspects. This approach ensures that the derived codes are not distorted by prior theoretical considerations [39, 40]. In contrast some scholars object to this view by mentioning the importance of ex ante considerations [41, 42]. Miles and Huberman (1984) state that it is not possible to conduct any qualitative analysis without having a conceptual framework that emerges from considerations prior to the data analysis [43]. Therefore, coding techniques either can be used on textual data without any prior theoretical considerations (e. g. grounded theory [40]) or with a theory orientated conceptual framework as a starting point (e. g. deductive qualitative content analysis [30]).

Besides theoretical considerations the researcher has to consider specific guidelines while the data gathering an coding process to ensure a solid data basis for further analysis (see [31, 36, 44]). Negations or intonations are just two sources for biases that occur within the coding processes and can be handled with separate coding iterations.

Although coding can be done manually, software tools (e.g. *NVivo*, *Atlas.Ti, HyperResearch, MaxQDA*) are used to support the researcher in the coding process. Besides the main functionality of attaching codes to text segments the tools provide options for quantification and visualization of results. These different types of illustrations and representations are valuable sources for data interpretation.

2.3 Quality Criteria and Shortcomings in Qualitative Research

The essential and conventional research quality criteria validity and reliability also apply for qualitative research. Notwithstanding the criteria do have other characteristics when used in this domain [45].

Validity describes the quality and information value an approach is able to measure in regard to what it claims to measure. To exclude interfering influences quantitative methods are highly standardized. One can see this standardization in an understated form in qualitative research as well. Especially the approaches for data collection and processing are equipped with rules to ensure a standardized process (e.g. corpus construction [46], coding in grounded theory [27]). But as Flick (2014) mentions the strength of approaches in the qualitative domain often lies in their flexibility [47]. In addition, he indicates that communicational effects are not controllable in a reasonable way. Besides standardization, the credibility and accuracy in qualitative research can be ensured via communicative validation [45, 48]. This can be performed in form of direct member checks that require a validation of the involved people. For example, an interview partner validates the transcript to ensure its correctness. Additionally, the accuracy of the qualitative approach can be confirmed via peer validation with other researchers or experts in the field to ensure that the right methodology is used.
Reliability stands for the robustness of findings and the overall consistency of an approach which is repeated with consistent conditions. In qualitative research identical results not always represent reliable findings [49]. E.g. identical responses in interviews are not a decent indicator for reliability but point to prepared answers which result in falsified findings. For reliable results in qualitative research the context in which the data is collected and the analyses are performed must be described in very detail. This intersubjective tractability helps to understand each part of the study and should lead to other researches drawing consistent conclusions [50, 51].

Besides reliability and validity additional quality criteria are important that are addressed with the proposed approach. Lincoln and Guba introduce credibility as a construct that ensures the internal validity of qualitative research [45]. They emphasize the trustworthiness of findings and recommend approaches like triangulation and negative case analysis. In addition, the intersubjective traceability should be given to ensure that the results and interpretations are objective and confirmable [42]. Although there exist guidelines on how to meet this quality requirement, critics blame the domain for its strong subjectivity when it comes to interpretation of analyzed data [52]. Since the researcher is the one who is collecting, processing, analyzing and interpreting the data some degree of bias is introduced because of his subjectivity [53]. To control this bias, the research progress should be enhanced with techniques that enable a standardized evaluation and interpretation of qualitative data.

3 An Approach to Generate a Deeper Insight into the Relevance of Coded Qualitative Data

Regarding the lack of techniques for objective interpretation and the need to meet the quality criteria credibility and transparency an approach is provided to gather hidden information from coded qualitative data. The presented technique embraces the numerical structure of coded data which results from the assignment of text passages to codes. With the technique the relevance of codes (i.e. an empirical generalization of relevant information in regard to the research question) can be identified. Based on an information retrieval method an index is calculated to compare codes concerning their relevance with regard to the underlying research objective.

3.1 Relevance of Codes Based on Frequency and Appearance

In order to generate conclusions from the coded data it is helpful to compare the identified codes regarding their importance (e.g. [4]). Therefore, a value is necessary to rank each code in respect to its relevance. To calculate this index a standardized procedure is needed that identifies the importance based on the given coding of raw data. In information science the concept of information retrieval describes the process of identifying relevant information resources in collections based on given information needs [54, 55]. As a measure of relevance the two components term frequency and inverse document frequency are used [56, 57]. The first measure covers the frequency of a term by counting its occurrence within the information source. The second measure extends the term frequency by considering not a single information source but all existing sources. Therefore, the inverse document frequency counts the existence of the term in question over all considered information sources. As Robertson (2004) points out, multiplying the two constructs generates "[…] extraordinarily robust and difficult to beat […]" results [56]. Considering this, I adapted the measurement and developed an approach to evaluate the relevance of a code.

3.1.1 Concept

I modified the measurement to make it suitable in the context of qualitative data. This is possible because in the coding process each source of information (e.g. interview transcript) is screened multiple times to identify relevant information (see subsection 2.2). Hence a code can be seen as a term of question in the information retrieval domain. Transferring the mechanism, one can assume, that on the one hand a code is relevant the more often it is mentioned within a single information source. On the other hand, a code is important the more information sources contain the respective code. The following table lists the used mathematical notations and describes their meaning to measure the relevance of codes.

Table 2. Mathematical Notations and Descriptions

Notation	Description
n	Number of information sources
m	Number of codes
i	Index of information sources, $i \in \{1, \dots, n\}$
c	Index of codes, $c \in \{1, \dots, m\}$
f_c	Code-frequency of code *c* within all information sources, $f_c \in \mathbb{N}_0$
f_{ci}	Code-frequency of code *c* within information source *i*, $f_{ci} \in \mathbb{N}_0$
a_{ci}	Binary variable representing the appearance of code *c* in source *i*
$CARFI_c$	Index indicating the relevance of code *c*, $CARFI_c \in \mathbb{Q}^+$
AI_c	Additional information for the relevance of a code *c*, $AI_c \in \mathbb{Q}^+$
w	Wight for additive linkage of AI_c, $w \in [0,1]$
$ECARFI_c$	Extended index indicating the relevance of a code *c* based on $CARFI_c$ and AI_c, $ECARFI_c \in [0,1]$

Regarding the code-frequency (f_{ci}) each information source should have the same weighting (relative code-frequency). Therefore, the number of coded text segments for a particular code *c* in an information source *i* has to be standardized by the total number of codes mentioned in the information source *i*. With regard to Namey et al. (2008) the source-frequency is taken into account as a second measure [58]. It represents the number of information sources which mention the particular code *c* and is calculated with the binary variable a_{ci}:

$$a_{ci} = \begin{cases} 1, if\ code\ c\ is\ mentioned\ in\ source\ i \\ 0, otherwise \end{cases}$$

The following formula represents the relevance of a code *c* containing the two factors relative code-frequency and source-frequency, where c' represents the total number of codes mentioned in information source *i*:

$$CARFI_c = \sum_{i=1}^{n} \frac{f_{ci}}{\sum_{c'=1}^{m} f_{c'i}} \sum_{i=1}^{n} a_{ci} \quad (1)$$

Regarding $CARFI_c$ it is assumed that codes with a high relative code-frequency combined with a high source-frequency matter more and hence have a higher relevance. This is expected because the more often a code appears in an information source relative to the other codes the more important the aspect is within this source. Because of this the relative code-frequency represents an indicator for the relevance of an aspect within an information source. But being coded more often than other codes within a source of information, does not implicate a high relevance. E. g. one must assume that a code which is only existing in one single source has less relevance than a code which is mentioned in all sources. Therefore, the source-frequency is additionally taken into consideration. It counts in how many different sources the code appears. Hence this measure gives an insight into the topic related importance of a code regarding the context of information sources. With the multiplicative connection of the two measures codes are considered as relevant if both the source-frequency and the relative code-frequency result in high scores.

3.1.2 Extended Concept to Include Additional Information

In addition to the absolute and relative frequency other dimensions like intonations can be an indicator for relevance and hence should be considered in the approach. This additional information about the relevance of a code c AI_c can be included with an additive connection of constructs (see equation 2). The additive link is realized with a specific weight w to integrate the additional information in the relevance ranking.

$$ECARFI_c = (1-w)\left(\frac{CARFI_c}{\sum_{c'=1}^{m} CARFI_{c'}}\right) + w\ \left(\frac{AI_c}{\sum_{c'=1}^{m} AI_{c'}}\right) \tag{2}$$

To ensure that the two measures in equation 2 can be combined, both terms of the sum need to be standardized (see denominators of the two terms). Additionally, the sum of weights must equal 1 which leads to the standardized relevance of a code (first term) being weighted with $1-w$ where w stands for the weight with which the additional information is considered. $ECARFI_c$ itself results in a number between 0 and 1 due to the standardization of the related terms.

The same linkage can also be done when information about validation or rankings of codes is available from other information sources (e.g. expert interviews to validate or rank codes). If more than one additional information should be included this can be done with an adaption of weights, where each additional information is linked additive with its own weight where the sum of weights for each term equals one.

If no additional information AI_c is available, $CARFI_c$ as well as $ECARFI_c$ can be used to indicate the relevance of codes. While $CARFI_c$ results in an absolute number $ECARFI_c$ calculates to a relative number, because no additional information AI_c is available ($AI_c = 0, w = 0$). If no term of the sum is emphasized the weights for each term of the sum equal 1 divided by the total number of terms. For example, if the two terms in equation 2 should be considered equally, then $w = 1/2$.

3.1.3 Example with Textual Data from Interviews

The technique has been tested with real data from interviews with entrepreneurship experts to identify success indicators for IT-startups [59]. Eleven interview transcripts were coded on the bases of the methodology proposed from Steigleder (2008). The theory orientated content coding technique results in 22 codes with each of them representing a separate success indicator. The data set in table 3 represents the results of the coding process of textual data from eleven interviews which results in 22 codes. Each code stands for a success indicator for IT-startups.

Table 3. Data Set of Coded Qualitative Data from Interviews

		code c																					
	i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
interview i	1	2	0	4	2	2	1	3	2	2	0	2	1	1	0	0	1	0	0	0	0	0	1
	2	3	0	1	1	0	2	0	1	2	0	3	1	0	0	0	5	7	1	4	1	2	0
	3	5	0	0	1	3	0	2	0	0	1	0	1	1	7	0	0	0	0	0	3	1	0
	4	2	1	0	2	3	2	1	1	1	0	0	2	7	0	3	0	1	3	3	1	0	2
	5	0	2	2	2	1	1	0	1	0	1	2	2	0	0	1	5	4	5	3	3	3	7
	6	2	2	2	2	0	1	0	0	2	4	0	0	1	0	4	1	1	0	0	3	0	0
	7	1	1	0	2	1	1	0	0	3	0	0	0	1	0	2	1	4	0	1	0	0	1
	8	2	1	0	2	3	4	0	0	1	1	2	2	6	0	2	2	5	2	2	2	5	1
	9	3	0	0	0	1	2	1	1	3	0	1	2	0	1	0	1	1	3	3	2	3	1
	10	0	0	0	0	0	0	0	0	0	2	1	3	2	2	0	0	0	0	0	2	1	2
	11	1	0	0	2	0	0	2	3	0	0	2	0	0	0	1	1	0	1	0	1	1	1

The values given in the matrix represent the code-frequency of a particular code in an information source (f_{ci}). E. g. code 1 is coded two times in interview 1 which results in f_{11} equals two. The sum of one column results in f_c , e. g. code 1 is coded 21 times in total.

After the interview the experts have been directly asked about their top five factors for startup success. This data has been taken into account as additional information AI_c regarding the relevance for a code. In the example AI_c represents how often a success indicator, i.e. a code c, has been given as answer to this question. E. g. code 20 has six times been mentioned as top five indicator, see table 4.

By calculating the values of $ECARFI_c$ the two aspect ($CARFI_c$ and AI_c) both should be considered equally, i.e. both indicate success in the same manner. This results in $w = 0{,}5$. After applying the technique, the codes can be arranged in an ordered sequence which represents the relevance ranking for the analyzed codes and serves as a foundation for interpretation. The results are represented in table 4 and support the researcher in identifying the relevance and importance of success indicators regarding the underlying research question. Although *handling of market conditions* (rank 11, $ECARFI_{17}$ = 0,042) is the most commonly coded indicator (f_{17} = 23) it is ranked in the middle of the list. This results from a low relative code-frequency and a low value regarding the additional information ($AI_{17} = 1$). Besides that, *perseverance* (rank 8, $ECARFI_2$ = 0,054) with the lowest code-frequency (f_2 = 7) is ranked three spots above *handling of market conditions* because the additional information about the relevance (AI_2 = 5) is considered and indicates an important aspect for startup-success. This shows that experts do not accentuate the importance of being able to handle complex market conditions as much as they emphasize *perseverance* in the interviews.

Table 4. Ranking of Codes Based on *ECARFI*

Rank	Sucess indicator	c	f_c	$\sum_{i=1}^{n} a_{ci}$	$CARFI_c$	AI_c	$ECARFI_c$
1	Customer orientated problem solving	20	18	9	6,07	6	0,095
2	Sales competence and marketing power	1	21	9	6,95	4	0,082
3	Seed-customer and technology partner	16	17	8	4,27	5	0,074
4	Feedback driven product developement	21	16	7	3,56	5	0,070
5	Prototype orientated product developement	22	16	8	4,48	4	0,066
6	Team composition	12	14	8	4,21	4	0,064
7	Industry specific competence	5	14	7	3,26	4	0,058
8	Perseverance	2	7	5	1,14	5	0,054
9	Scalability and market ability	9	14	7	3,74	3	0,052
10	Staff management skills	8	9	6	2,31	3	0,043
11	Handling of market conditions	17	23	7	5,04	1	0,042
12	Industry specific financing	13	19	7	4,49	1	0,039
13	Entrepreneurship and professional experience	4	16	9	5,48	0	0,036
14	Political and regulartory business environment	18	15	6	2,62	2	0,036
15	ICT competence	6	14	8	3,44	1	0,032
16	Conversion capability and speed	7	9	5	1,97	2	0,031
17	Proof of feasability and verification	10	9	5	1,89	2	0,031
18	Industry specifc norms and requirements	19	16	6	2,82	1	0,028
19	Business model felxibility and independence	11	13	7	3,41	0	0,022
20	R&D cooperations	15	13	6	2,88	0	0,019
21	Value orientated thinking	3	9	4	1,28	1	0,018
22	Accelerator or incubator program	14	10	3	1,34	0	0,009

Besides that, one can see that codes with the same source-frequency can be brought to an order. E. g. the *sales competence and marketing power* ($ECARFI_1$ = 0,082) as well as the *entrepreneurship and professional experience* ($ECARFI_4$ = 0,036) were both mentioned from 9 different experts but are ranked on second and 13th place with a difference of 0.036. This results from the fact that experts emphasize the sales skills of the startup team within their interviews and mention it as top 4 success factor. Besides identifying the indicators of IT-startup success, the ranking makes it possible to give recommendations about the relative importance of each aspect which cannot be revealed from a list of simple code frequencies.

4 Limitations and Future Research

Although the approach generates stable results it is subject to some limitations. First, the quality of the results generated can only be as good as the quality of the underlying coding process. Therefore, using the techniques does neither understate the need for a solid data collection nor does it replace a robust coding of the qualitative data. It has to be seen as an additional measure which helps to detect hidden information in coded datasets. Second, although the approach generates neutral and objective results, the researcher still interprets them from a subjective perspective that can result in biases.

Third, it is important to understand that the results only provide an indication for relevance but cannot generate findings that are valid in general. This is due to the structure of qualitative research which aims to understand complex phenomena regarding a specific topic of interest. Fourth, frequency is just one approach to address relevance in an objective manner but does not fit for every qualitative investigation. To grant stable findings it is important to combine the proposed technique with other approaches to judge relevance (e.g. context a code comes from or is related to, verbal traces and actual speech such as laughter). In addition to that supplementary mechanisms can be applied to make the input more robust (e. g. case contrasting proposed by Flick (2014) diminishes the impact of similar sources). Fifth, the approach views each source as equal and includes each code mentioned in a source based on the total amount of codes within this source. Therefore, codes mentioned in sources with a lot of codes are less emphasized then codes mentioned in sources with few codes. This behavior must be considered when using the technique.

Future research should investigate the fit of the presented approaches for qualitative research in more detail. To check the robustness of the ranking, one could use a linear ordering technique which examines the ranking based on additional data sources indicating the ranking. Furthermore, intentionally a small data set was applied to introduce the approach in a simple and understanding manner. Although the approach has been used with large data samples in its original discipline, the next step is the investigation of the behavior with different dimensions of data size in qualitative terms.

5 Conclusion

This paper introduces a new qualitative research approach which expands the existing repertoire of qualitative data analysis methods. Therefore, I draw on an existing approach from the information science discipline to bring new insights to current qualitative methods in IS. The technique is adopted and enhanced to be used on data which results from any kind of qualitative coding process. Hence, the approach can be performed independently of the underlying methodology and does not intend to replace existing procedures but aims to extend them.
By developing the practice, I focused on answering questions about the relevance of codes. The proposed technique results in an index representing the relevance or importance of a code relative to other codes. The generated results can serve as an additional basis for analysis and interpretation of qualitative data sources. Until now analysis was only based on the system of codes and their assigned textual passages. With the additional findings from my approach a deeper insight can be generated that supports the researcher in detecting hidden information in the coded data. Regarding the intersubjective tractability and credibility the standardized measure enables replicable and transparent findings generated from qualitative data. The formalization of the approach makes it obvious how the results revealed. With this, one can counteract the quality issues tractability and credibility on the one hand and fulfill the principle of transparency on the other.

The technique addresses and complements many existing and future qualitative and mixed method studies based on coding techniques in the IS domain. Especially when it comes to large empirical data samples in terms of big data the full potential of the approach can be exploited. As the complexity increases with a growing number of sources the mechanism can support analyzing and interpreting “big qualitative data”.

References

1. Sarker, S., Xiao, X., Tanya, B.: Towards an Anatomy of “Successful” Qualitiative Research Manuscripts in IS: A Critical Review and Some Recommendations. Thirty Third Int. Conf. Inf. Syst. (ICIS 2012). 1–21 (2012).
2. Chua, C.E.H., Yeow, A.Y.K.: Artifacts, actors, and interactions in the cross-project coordination practices of open-source communities. J. Assoc. Inf. Syst. 11, 838–867 (2010).
3. Davison, R.M., Vogel, D.R.: Group Support Systems in Hong Kong: An Action Research Project. Inf. Syst. J. 10, 3–20 (2000).
4. Vogelsang, K., Steinhüser, M., Hoppe, U.: A Qualitative Approach to Examine Technology Acceptance. Int. Conf. Inf. Syst. (ICIS 2013). 234–245 (2013).
5. Urquhart, C., Lehmann, H., Myers, M.D.: Putting the “theory” back into grounded theory: guidelines for grounded theory studies in information systems. Inf. Syst. J. 20, 357–381 (2010).
6. Klein, H.K., Myers, M.D.: A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. MIS Q. 23, 67–93 (1999).
7. Schultze, U.: A Confessional Account of an Ethnography about Knowledge Work. MIS Q. 24, 3–41 (2000).
8. Conboy, K., Fitzgerald, G., Mathiassen, L.: Qualitative Methods Research in Information Systems: Motivations, Themes, and Contributions. Eur. J. Inf. Syst. 21, 113–118 (2012).
9. Marton, A.: Purposive Selection and the Quality of Qualitative IS Research. Thirty Fourth Int. Conf. Inf. Syst. 1–19 (2013).
10. Sarker, S., Xiao, X., Beaulieu, T.: Guest Editorial: Qualitative Studies in Information Systems: A Critical Review and Some Guiding Principles. Manag. Inf. Syst. Q. 37, (2013).
11. Asatiani, A.: Why Cloud? - A Review of Cloud Adoption Determinants in Organizations. Twenty Third Eur. Conf. Inf. Syst. (ECIS 2015). 1–17 (2015).
12. Lang, M., Wiesche, M., Krcmar, H.: What are the most important Criteria for Cloud Service Provider Selection? A Delphi Study. Twenty Fourth Eur. Conf. Inf. Syst. (ECIS 2016). 1–17 (2016).
13. Myers, M.D.: Interpretive Research Methods in Information Systems. In: Mingers, J. and Stowell, F. (eds.) Information Systems: An Emerging Discipline. pp. 239–266. McGraw Hill, London (1997).
14. Venkatesh, V., Brown, S.A., Bala, H.: Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems. Manag. Inf. Syst. Q. 37, 21–54 (2013).
15. Johnson, R.B., Turner, L.A.: Data Collection Strategies in Mixed Methods Research. In: Tashakkori, A. and Teddlie, C. (eds.) Handbook of Mixed

Methods in Social & Behavioral Research. pp. 297–319. Sage Publications, Thousand Oaks, CA (2003).
16. Wilde, T., Hess, T.: Forschungsmethoden der Wirtschaftsinformatik Eine empirische Untersuchung. Wirtschaftsinformatik. 49, 280–287 (2007).
17. Susman, G.I., Evered, R.D.: An Assessment of the Scientific Merits of Action Research. Adm. Sci. Q. 23, 582–603 (1978).
18. Baskerville, R.L., Wood-Harper, A.T.: Diversity in information systems action research methods. Eur. J. Inf. Syst. 7, 90–107 (1998).
19. Ytterstad, P., Akselsen, S., Svendsen, G., Watson, R.T.: Teledemocracy: Using Information Technology to Enhance Political Work. MISQ Discov. 1, (1996).
20. Yin, R.K.: The Case Study as a Serious Research Strategy. Knowledge. 3, 97–114 (1981).
21. Yin, R.K.: Case Study Research: Design and Methods. Sage Publications, Newbury Park (2003).
22. Thomas, G.: A Typology for the Case Study in Social Science Following a Review of Definition, Discourse, and Structure. Qual. Inq. 17, 511–521 (2011).
23. Alavi, M., Carlson, P.: A review of MIS research and disciplinary development. J. Manag. Inf. Syst. 8, 45–62 (1992).
24. Lewis, I.M.: Social Anthropology in Perspective. Cambridge University Press, Cambridge (1985).
25. Wynn, E.: Office conversation as an Information Medium, (1979).
26. Baskerville, R.L., Myers, M.D.: Design ethnography in information systems. Inf. Syst. J. 25, 23–46 (2015).
27. Strauss, A., Corbin, J.: Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Sage Publications, Newbury Park, CA (1990).
28. Martin, P.Y., Turner, B.A.: Grounded Theory and Organizational Research. J. Appl. Behav. Sci. 22, 141–157 (1986).
29. Mayring, P.: Qualitative Inhaltsanalyse: Grundlagen und Techniken. Beltz, Weinheim (2015).
30. Steigleder, S.: Die strukturierende qualitative Inhaltsanalyse im Praxistest: eine konstruktiv kritische Studie zur Auswertungsmethodik von Philipp Mayring, (2008).
31. Saldaña, J.: The coding manual for qualitative researchers. Sage Publications, Los Angeles (2016).
32. Taylor, C.: Hermeneutics and Politics. In: Connerton, P. (ed.) Critical Sociology: Selected Readings. pp. 153–193. Penguin Books, Harmondsworth (1976).
33. Gadamer, H.-G.: Philosophical Hermeneutics. University of California Press, California (1976).
34. Krippendorff, K.: Content analysis : An introduction to its methodology. Sage Publications, Beverly Hills (1980).
35. Howcroft, D., Trauth, E.M.: Handbook of Critical Information Systems Research: Theory and Application, (2005).
36. Gläser, J., Laudel, G.: Life with and without coding: Two methods for early-stage data analysis in qualitative research aiming at causal explanations. Forum Qual. Sozialforsch. 14, (2013).
37. Kelle, U.: Theory Building in Qualitative Research and Computer Programs for

the Management of Textual Data. Sociol. Res. 2, (1997).
38. Coffey, A., Atkinson, P.: Making sense of qualitative data: Complementary research strategies. (1996).
39. Meinefeld, W.: Ex-ante-Hypothesen in der Qualitativen Sozialforschung: zwischen “fehl am Platz” und “unverzichtbar”. Z. Soziol. 26, 22–34 (1997).
40. Glaser, B., Strauss, A.: The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine Publishing Company, Chicago (1967).
41. Auer-Srnka, K.J.: Hypothesen und Vorwissen in der qualitativen Marktforschung. In: Qualitative Marktforschung. pp. 159–172. Springer (2009).
42. Bogner, A., Littig, B., Menz, W.: Interviews mit Experten: Eine praxisorientierte Einführung. Springer Fachmedien, Wiesbaden (2014).
43. Miles, M.B., Huberman, A.M.: Qualitative Data Analysis: A Sourcebook of New Methods. Sage Publications, Newbury Park, CA (1984).
44. Campbell, J.L., Quincy, C., Osserman, J., Pedersen, O.K.: Coding In-depth Semistructured Interviews: Problems of Unitization and Intercoder Reliability and Agreement. Sociol. Methods Res. 42, 294–320 (2013).
45. Lincoln, Y.Y., Guba, E.G.: Naturalistic Inquiry. Sage, Beverly Hills (1985).
46. Bauer, M.W., Gaskell, G.: Qualitative researching with text, image and sound: A practical handbook for social research. Sage (2000).
47. Flick, U.: Gütekriterien qualitativer Sozialforschung. In: Baur, N. and Blasius, J. (eds.) Handbuch Methoden der empirischen Sozialforschung. pp. 411–423. Springer, Wiesbaden (2014).
48. Bryman, A.: Research Methods and Organization Studies. Unwin Hyman, London (1989).
49. Flick, U.: An introduction to qualitative research. Sage, Los Angeles (2014).
50. Kruse, J.: Qualitative Interviewforschung : Ein integrativer Ansatz, (2015).
51. Steinke, I.: Kriterien qualitativer Forschung. Juventa Verlag, Weinheim München (1999).
52. Wrona, T., Gunnesch, M.: The One Who Sees More is More Right: How Theory Enhances the “Repertoire to Interpret” in Qualitative Case Study Research. J. Bus. Econ. 1–27 (2015).
53. Zahedi, F.M., Van Pelt, W.: Web Documents ’ Cultural and Femininity Masculinity. J. Manag. Inf. Syst. 23, 87–128 (2006).
54. Belkin, N.J., Croft, W.B.: Information Filtering and Information Retrieval: Two Sides of the Same Coin? Commun. ACM. 35, 29–38 (1992).
55. Salton, G., McGill, M.J.: Introduction to Modern Information Retrieval. (1986).
56. Robertson, S.: Understanding inverse document frequency: on theoretical arguments for IDF. J. Doc. 60, 503–520 (2004).
57. Wu, H.C., Luk, R.W.P., Wong, K.F., Kwok, K.L.: Interpreting TF-IDF term weights as making relevance decisions. ACM Trans. Inf. Syst. 26, 1–37 (2008).
58. Namey, E., Guest, G., Thairu, L., Johnson, L.: Data Reduction Techniques for Large Qualitative Data Sets. Handb. Team-based Qual. Res. 137–163 (2008).
59. Keller, A.: Zum Erfolg von IKT-Start-ups in der deutschen Elektrizitätswirtschaft - Eine explorative Studie auf Basis von Experteninterviews. Verlag Dr. Kovač, Hamburg (2016).

On the Research Paradigms and Research Methods Employed in the BISE Journal – A Ten-Year Update

Thomas Friedrich, Sebastian Schlauderer, Julian Weidinger, and Maximilian Raab

University of Bamberg, Chair of Industrial Information Systems, Bamberg, Germany
{thomas.friedrich,sebastian.schlauderer,julian.weidinger,maximilian.raab}@uni-bamberg.de

Abstract. In the past decade, there has been an intense discussion in the German-speaking IS research community about the different paradigmatic orientations and the methodological diversity. With the work at hand, we present an overview of the research paradigms and research methods that have been used in the German-speaking IS community, represented by the BISE journal, over the past decade. After analyzing 169 research articles, we contribute an overall picture of research paradigms and research methods, a detailed picture of research methods per research paradigm, and a picture of historical trends in research paradigms and research methods. A comparison with previous studies reveals the unique profile of the German-speaking IS community. However, our results also indicate that there might be a shift in the use of research paradigms and research methods. Hence, this study provides a useful basis for future discussions about the positioning and the challenges of the community.

Keywords: IS Research, Research Paradigms, Research Methods, Meta-Analysis.

1 Introduction

In the past decade, there have been several debates in the German-speaking Information Systems (IS) research community on how the field should develop and position as a scientific discipline [1-3]. Especially the different paradigmatic views between behavior science research and design science research have led to extensive discussions [4-7]. In general, the behavior science research paradigm is concerned with the development and justification of theories that explain human or organizational behavior, while the design science research paradigm intends to create novel IT artifacts to solve organizational problems [8], [9]. Since the establishment of the IS discipline, design-oriented research has been the dominating research paradigm in many European countries, especially in the German-speaking countries where engineering disciplines have a strong position [10], [11]. In the international IS literature, the situation has been quite opposite because most articles published in top IS journals follow the behavior science paradigm [12], [13]. Given the increased pressure to publish in high-ranked international IS journals [14], researchers in the German-speaking IS community have thus been

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Friedrich, T.; Schlauderer S.; Weidinger, J.; Raab, M. (2017): On the Research Paradigms and Research Methods Employed in the BISE Journal - A Ten-Year Update, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1111-1125

concerned that design-oriented research approaches will be gradually displaced by behavior-oriented research approaches [15]. This would not only lead to a loss of identity of the German-speaking IS community but also reduce the methodological diversity which, however, is considered as critical to the relevance of the discipline [15-17].

As an interdisciplinary research field, IS draws upon a wide range of disciplines, such as informatics, economics, engineering, psychology, sociology, or mathematics [18], [19]. Consequently, a variety of research methods has been proposed and used to generate knowledge about the development, implementation, and use of information systems [10], [12], [20]. Meanwhile, several studies exist that examine the research methods and/or research paradigms of the IS discipline [e.g., 12], [13], [20], [21]. However, most of these studies target the international (predominantly the North American) IS community, while the German-speaking IS community has received little attention so far [5]. Well-known and frequently cited in the German-speaking IS community is the study by Wilde and Hess [10], who examined the methodological profile of the IS discipline of the German-speaking countries by examining research articles that had been published between the years 1996 and 2006 in the former German IS journal "WIRTSCHAFTSINFORMATIK". In 2014, the journal has been renamed to Business & Information Systems Engineering (BISE) journal and the publication language has been changed to English [22].

Given the existing debates about the positioning and the challenges of the German-speaking IS community, it becomes important to derive a clear understanding of the paradigms and the research methods that are currently being used in the community [5], [15]. Moreover, when considering the study by Wilde and Hess [10], it is interesting to see if the paradigms and research methods in the community have changed throughout the past decade. In this way, researchers can gain a better understanding of their own discipline and its underlying methodological variety. Such an understanding can moreover serve as basis for a well-informed discussion about the future development of a research community. Therefore, the aim of this paper is to provide a structured and up to date overview of the paradigmatic and methodological profile of the German-speaking IS community. In line with the study by Wilde and Hess [10], the German-speaking IS community in this paper is represented by research articles that have been published in the BISE journal, which describes itself as the "flagship journal of the German-language Information Systems community" [23]. In so doing, we remain able to compare our results with Wilde and Hess [10], which allows us to position this paper as a ten-year update. In particular, we address the following research questions:

1. What research paradigms and research methods have been used in the BISE journal in the past decade?
2. What are the most popular research methods per research paradigm in the BISE journal in the past decade?
3. Have the research paradigms and research methods in the BISE journal changed throughout the past decade?

To answer these research questions, we conducted a systematic meta-analysis on the research articles that have been published in the BISE journal in the last ten years. Research paradigms and research methods were analyzed, categorized, and counted per article. Furthermore, the results were compared with previous studies.

The remainder of the paper is structured as follows. In the next section, we describe the theoretical background. In section 3, we illustrate our research procedure. The results of our analysis are presented in section 4. In section 5, we discuss our results in comparison with other studies, followed by the implications and limitations. In section 6, we conclude by summarizing our results.

2 Theoretical Background

In this section, we provide background information on the research paradigms and the research methods that are used in the IS discipline.

2.1 Research Paradigms in the IS Discipline

According to Hirschheim and Klein [24, p. 1201], a *research paradigm* is "the most fundamental set of assumptions adopted by a professional community that allows its members to share similar perceptions and engage in commonly shared practices". It consists of assumptions about knowledge and how to acquire it, and about the physical and social world. In the IS literature, a variety of research paradigms, as well as classifications of research paradigms, can be found. For instance, Chen and Hirschheim [12], Weber [25], as well as Fitzgerald and Howcroft [26] distinct between a positivist and an interpretive research paradigm, while Iivari [27] as well as Goles and Hirschheim [28] distinct between a functionalist, an interpretivist, a radical humanist, and a radical structuralist paradigm. In these studies, the term research paradigm is used to distinct between different epistemological and ontological positions. However, the term research paradigm has also been established in the IS literature to refer to a problem-oriented process that consists of two phases, problem understanding and problem solving [8], [9]. According to Hevner, et al. [8, p. 76], the *behavior science research paradigm* is understood as a problem understanding paradigm that "seeks to develop and justify theories (i.e., principles and laws) that explain or predict organizational and human phenomena surrounding the analysis, design, implementation, management, and use of information systems. In contrast, the *design science paradigm* is understood as a problem-solving paradigm that "seeks to create innovations that define the ideas, practices, technical capabilities, and products through which the analysis, design, implementation, management, and use of information systems can be effectively and efficiently accomplished" [8, p. 76].

In this paper, we focus on the distinction between behavior science research and design science research because both research paradigms are frequently discussed in the literature when it comes to the positioning and development of the German-speaking IS community [4], [6], [7].

2.2 Research Methods in the IS Discipline

In general, a *research method* can be understood as a set of activities that must be undertaken to conduct research [29]. It is a "well-defined sequence of elementary operations which permits the achievement of certain outcomes if executed correctly" [30, p. 165]. Given the diverse nature of the IS discipline and the different paradigmatic positions, researchers have employed a wide range of research methods to study the design, implementation, and use of information systems. Consequently, various classifications of research methods can be found in the international IS literature [12], [13], [20], [31], [32]. Well-known and frequently-cited in the German-speaking IS literature is the classification of research methods developed by Wilde and Hess [10]. The classification scheme consists of the following research methods: *formal-*, *conceptual-*, and *argumentative-deductive analysis*, *simulation*, *reference modeling*, *action research*, *prototyping*, *ethnography*, *case study*, *grounded theory*, *qualitative/quantitative cross-section analysis*, and *laboratory/field experiment* (detailed descriptions of these methods can be found in Wilde and Hess [10]). We use the classification of research methods developed by Wilde and Hess [10] in this paper because it is well-grounded and adapted to the German-speaking IS literature. Moreover, it covers research methods from the behavior science as well as the design science research paradigm. In contrast, classification schemes proposed in the international IS literature typically do not cover design-oriented research methods (e.g., argumentative-deductive analysis, prototyping, or reference modeling) because these classifications have been developed on studies that follow the behavior science research paradigm [12], [13], [20], [31], [32].

3 Research Procedure

Our research procedure follows the meta-analysis approach of Palvia, et al. [20] and Palvia, et al. [21], who investigated the use of research methods and topical trends in the international IS literature. Note that this approach has also been labeled as descriptive review, which is a particular type of a literature review that introduces some quantification (e.g., through a frequency analysis) in order to verify a particular proposition or to reveal an interpretable pattern [33]. Following Palvia, et al. [20] and Palvia, et al. [21], our research procedure consisted of three stages: (i) data selection, (ii) data classification, and (iii) data analysis.

In the data selection stage, we selected all research articles published in the BISE journal from 2007 (volume 49, issue 1) to 2016 (volume 58, issue 4). The ten year period was chosen with reference to the study by Wilde and Hess [10], who analyzed research articles published in the BISE journal between 1996 and 2006. In line with Wilde and Hess [10], we selected only research articles that have been published in the section "Research Paper" of the BISE journal. Articles published in this section provide complete research results and substantial contributions to the literature. Articles published in other sections, such as "Catchword", "Interview", "Research Note", or "State of the Art", were excluded. Furthermore, the research articles published in volume 51, issue 1, 2009 of the BISE journal were excluded because this issue only includes former best paper articles and some retrospective views on the history of the IS field. After this stage, our final data sample consisted of 169 research articles.

In the data classification stage, we systematically analyzed and categorized the contents of each research article with respect to its research paradigm and its research methods. Each research article was carefully read and categorized by three research assistants independently. The results were then compared in a roundtable session and variations in the results were discussed until a consensus was reached. To assess the inter-coder reliability [34], a 10% random sample of articles was separately categorized by a fourth research assistant, resulting in an agreement of 94% (values of 80% or higher can be considered as acceptable, cf. [34]). According to our research questions, each article was categorized according to the both paradigmatic orientations, behavior science research and design science research (cf. section 2). In case the research paradigm has not been mentioned in an article, we followed the guidelines proposed by [8]. Articles addressing the design and evaluation of IT artifacts were assigned to the design science research paradigm, while articles addressing behavior-related aspects (e.g. use, adoption, or success of IT artifacts) were assigned to the behavior science research paradigm. In addition, we recorded the type of IT artifact that has been designed and/or evaluated in articles following the design science research paradigm. Categories for IT artifacts were constructs, models, methods, instantiations, and design theories [8], [9]. Furthermore, we analyzed the research methods that have been used in each article. For the categorization of research methods, we used the classification of research methods developed by Wilde and Hess [10] (cf. section 2). However, and in contrast to Wilde and Hess [10], we did not only focus on the primary research method of each research article (i.e., the research method that leads to the core contribution of the paper). Instead, we also captured all other research methods that were used and mentioned in an article. In this way, we were able to identify research articles that employed multiple research methods. The use of multiple research methods, for instance, is considered to be important in design science research in which IT artifacts are built and evaluated [8]. Therefore, we assigned research methods used in articles following the design science research paradigm to the build or to the evaluation process of the design science research project. Research methods used in articles following the behavior science paradigm were not assigned to specific stages because these behavior science research typically does not draw upon such a dichotomous research process.

In the data analysis stage, we analyzed the research articles based on the established categorization. The results of this stage are explained in the following section.

4 Data Analysis and Results

We structured the results of our analysis according to our research questions. First, we provide an overall picture of the research paradigms and the research methods that were used in the 169 analyzed research articles. Second, we provide a detailed picture of the use of research methods per research paradigm. Third, we provide a detailed picture of the historical trends in the use of research paradigms and research methods.

4.1 Overall Use of Research Paradigms and Research Methods

Figure 1 depicts the overall use of research paradigms (left side) and research methods (right side). With respect to the research paradigms, 65.1% (110) of the research articles follow the design science research paradigm, while 34.9% (59) of the research articles follow the behavior science research paradigm.

With respect to the use of research methods, the argumentative-deductive analysis (ADA) is used most frequently with 43.8%, followed by the case study (CS) with 30.8%. Other more frequently used research methods are the quantitative cross-section analysis (QNA) with 16.0%, the conceptual-deductive analysis (CDA) with 15.4%, and prototyping (PT) with 14.8%. All other research methods are used less frequently (<11%). Ethnography (ET) has not been used in any of the analyzed articles.

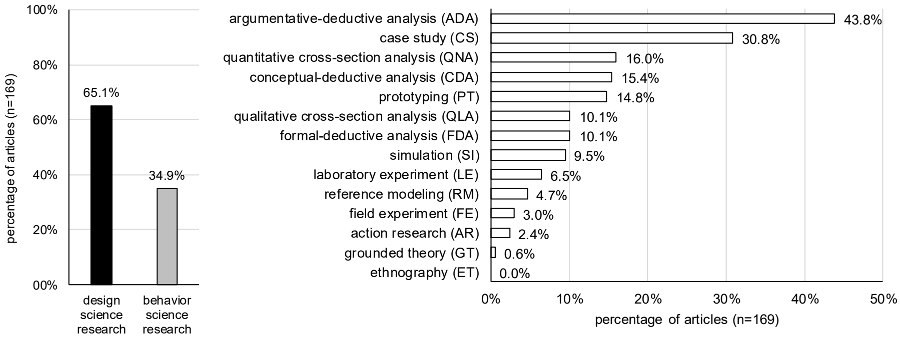


Figure 1. Overall use of research paradigms and research methods

4.2 Use of Research Methods per Research Paradigm

In the following, we first describe the results for the research articles following the design science research paradigm. Then, we describe the results for the research articles following the behavior science research paradigm.

Figure 2 shows the use of research methods of articles following the design science research paradigm according to the design science processes, build (left side) and evaluate (right side). In the building process, the most frequently used research method is the ADA with 47.3%, followed by the CDA with 22.7%. Other more frequently used methods to build IT artifacts are the formal-deductive analysis (FDA) with 14.5%, PT with 11.8%, and reference modeling (RM) with 7.3%. Qualitative cross-section analysis (QLA) with 1.8%, CS with 0.9%, and action research (AR) with 0.9% have rarely been used to build IT artifacts. Research methods that have not been used in the building process are excluded from the diagram. These methods are QNA, simulation (SI), laboratory experiment (LE), field experiment (FE), grounded theory (GT), and ET.

In the evaluation process, the most frequently used research method is the CS (32.7%), followed by the ADA (11.8%) and PT (10.9%). Other research methods that have been more frequently used to evaluate IT artifacts are the QLA (10.0%) and SI (10.0%). LE are used in 6.4%, FE in 3.6% of IT artifact evaluations. Rarely used in the evaluation process are AR (2.7%) and the QNA (1.8%). Research methods that have not been used in the evaluation process and that we excluded from the diagram are the CDA, the FDA, RM, GT, and ET.

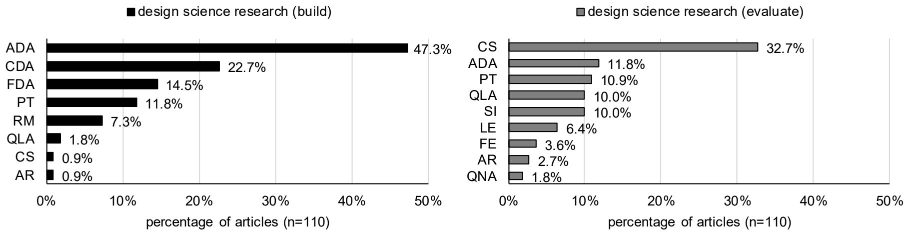


Figure 2. Use of research methods in design-oriented research articles

The separation between the building and the evaluation process enabled us also to investigate frequent combinations of research methods. The diagram on the left side of Figure 3 depicts how often research methods in the building and in the evaluation process have been used in combination. For instance, in 18 research articles, an ADA was used to build an IT artifact together with a CS to evaluate the IT artifact. In nine research articles, an IT artifact was built by using a CDA and the IT artifact was evaluated through a CS. In six research articles, an IT artifact was built by using a FDA and the IT artifact was evaluated through a SI.

The diagram on the right side of Figure 3 illustrates the number of research methods that have been used in design-oriented research articles. 24.5% of the design-oriented research articles only used one research method. In these articles, research methods were only used to build an IT artifact, but not to evaluate it. In 55.5% of the design-oriented research articles, two research methods were used to build and evaluate an IT artifact, while 18.2% of the articles used three research methods. Four research methods have been used in 1.8% of the design-oriented research articles.

method (build) \ method (evaluate)	ADA	SI	AR	PT	CS	QLA	QNA	LE	FE
FDA	0	6	0	0	5	0	0	0	0
CDA	5	2	0	4	9	1	0	2	0
ADA	7	1	1	6	18	6	1	2	3
RM	1	0	2	1	3	4	0	0	0
AR	0	0	0	0	0	1	0	0	0
PT	2	1	0	0	2	1	1	3	2
CS	0	0	0	0	0	1	0	0	0
QLA	1	0	0	0	0	0	0	0	0

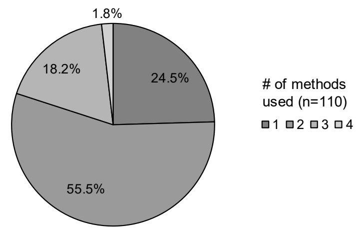
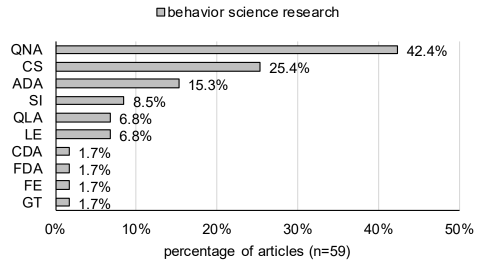


Figure 3. Combinations and number of research methods in design-oriented research articles

Considering the outcomes of the design-oriented research articles, the most frequent IT artifacts are methods (38.2%) and models (34.5%). Constructs have been build and/or evaluated in 17.3% of the articles. In 9.1% of the articles, the outcome represented instantiations. Design theories have been the outcome of 2.7% of the articles. The latter result reflects recent discussions in the literature that design science research should focus more on the development of design theories [35], [36].

The diagram on the left side of Figure 4 shows the use of research methods of articles following the behavior science research paradigm. In this context, the most frequently used research method is the QNA with 42.4%, followed by the CS with 25.4%. Other more frequently used research methods are the ADA (15.3%), SI (8.5%), the QLA (6.8%), and the LE (6.8%). CDA, FDA, FE, and GT have only been rarely used in these articles. Research methods that have not been used in the behavior-oriented research articles are PT, RM, AR, and ET. We excluded these methods from the diagram.



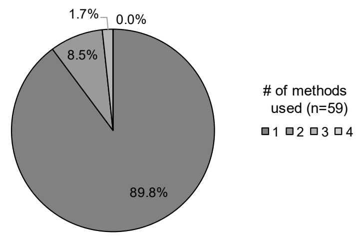


Figure 4. Use and number of research methods in behavior-oriented research articles

The diagram on the right side of Figure 4 depicts the number of research methods used in the behavior-oriented research articles. Most of these articles (89.8%) employ one research method. Two research methods are used in 8.5% of the articles, while three research methods are only used in 1.7% of the articles.

4.3 Historical Trends in Research Paradigms and Research Methods

To identify potential changes in the use of research paradigms and research methods, we analyzed the corresponding historical trends. Figure 5 illustrates the trend of research articles following the design science research paradigm vs. research articles following the behavior science research paradigm.

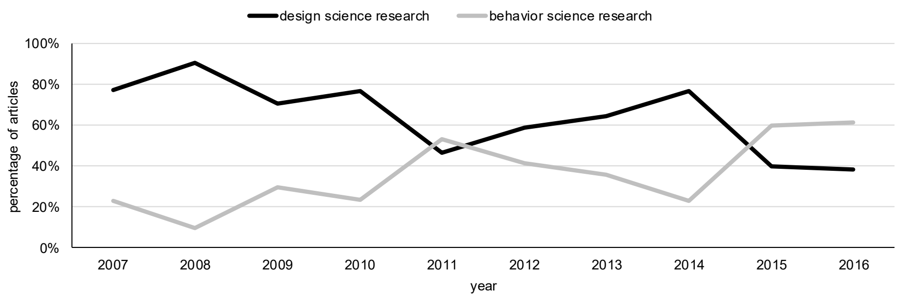


Figure 5. Use of research paradigms over the past decade

From 2007 to 2010, design-oriented research articles were in the majority in the BISE journal (on average, 78.7% design-oriented vs. 21.3% behavior-oriented). In 2011, the percentage of behavior-oriented articles was slightly above design-oriented articles with 53.3% and 46.7% respectively. Between 2012 and 2014, the percentage of design-oriented articles increased from 58.8% to 76.9%, while the percentage of behavior-oriented articles decreased from 41.2% to 23.1%. Since 2015, the trend has changed to, on average, 39.2% design-oriented articles and 60.8% behavior-oriented articles. Furthermore, we analyzed potential changes in the research methods over the past decade. Figure 6 illustrates the trend for the top 5 research methods.

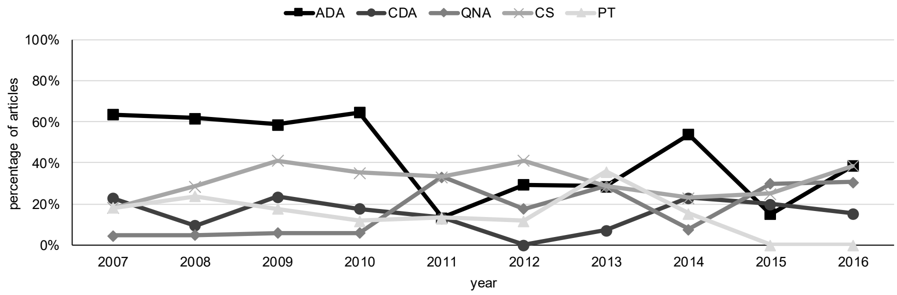


Figure 6. Use of research methods (overall top 5) over the past decade

Between 2007 and 2010, the ADA was the most frequently used research method (on average, 62.3%). In 2011, its use decreased to 13.3%, while between 2012 and 2014, its use increased to 53.9%. In 2015, its use again decreased to 15%, while in 2016, its use increased to 38.5%. Between 2007 and 2009, the use of CS increased from 18.8% to 41.2%. Since 2010, CS have been used, on average, in 33.3% of the articles. Between 2007 and 2010, the QNA has been used, on average, in 5.3% of the articles. Between 2011 and 2013, it was used, on average, in 26.5% of the articles, while in 2014, its use decreased to 7.7%. Between 2015 and 2016, the QNA has been used, on average, in 30.4% of the articles. Throughout 2007 and 2014, PT was used, on average, in 18.5% of the articles. However, since 2015, PT has not been employed in any of the analyzed articles. Between 2007 and 2016, the CDA has been used, on average, in 15.2% of the articles. In 2012, the CDA was not used in any article.

In addition, we analyzed the trend on the average number of research methods used per article over the past decade. Between 2007 and 2012, on average, 1.7 research methods were used per article. In 2013, the average number of methods increased to 2 methods per article. Between 2014 and 2015, the average number of methods decreased to 1.3 methods per article. In 2016, the average number of methods per article was 1.6.

5 Discussion

In the following subsections, we first compare our results with the study by Wilde and Hess [10] and Palvia, et al. [21], who examined the international IS literature. Then we discuss the implications and limitations of our study.

5.1 Comparison of Findings

Our results demonstrate that the majority (65.1%) of the 169 analyzed research articles follow the design science research paradigm, while 34.9% of the articles follow the behavior science research paradigm. When comparing our results with the study by Wilde and Hess [10], who analyzed research articles published in the BISE journal between 1996 and 2006, we notice a slight decrease in the percentage of design-oriented research articles. In the data sample of Wilde and Hess [10], the percentage of design-oriented and behavior-oriented research articles was 70% and 30% respectively. However, our results also reveal that since 2015, about 60% of the articles follow the behavior science research paradigm, while 40% of the articles follow the design science research paradigm. The results are in contrast to Wilde and Hess [10], who concluded that between 1996 and 2006 no indications can be found that lead to a shift in the research paradigms.

Figure 7 compares the overall use of research methods with the study by Wilde and Hess [10]. As suggested by Wilde and Hess [10], research methods are grouped according to their degree of formalization and according to the research paradigm in which these methods are typically applied. Note that the comparison only covers primary research methods (i.e., research methods that lead to the core contribution) since Wilde and Hess [10] only focused on the primary research methods (cf. section 3).

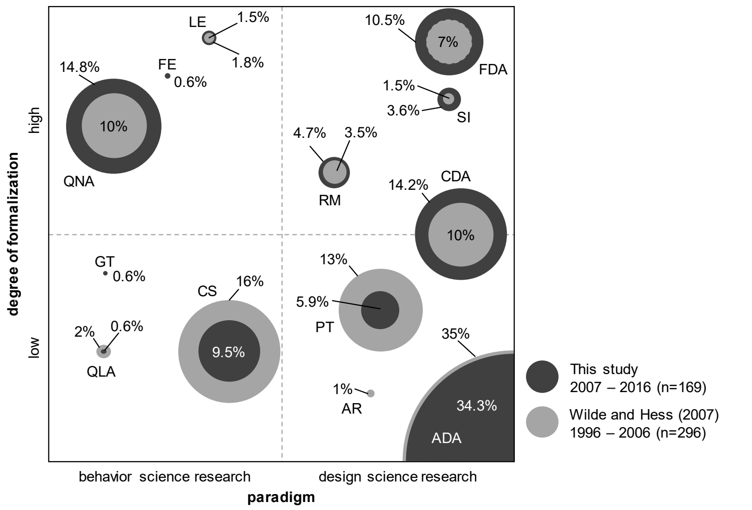


Figure 7. Use of research methods compared to Wilde and Hess [10] (only primary methods)

The results of the comparison illustrate that the ADA hast been the most frequently used method over the last twenty years (from 35% to now 34.3%). Moreover, we can observe an increased use of research methods with a higher degree of formalization, such as the QNA (from 10% to 14.8%), the CDA (from 10% to 14.2%), the FDA (from 7% to 10.5%), or RM (from 3.5% to 4.7%). In turn, the use of research methods with a lower degree of formalization decreased, such as the CS (from 16% to 9.5%), PT (from 13% to 5.9%), or the QLA (from 2% to 0.6%). In the study by Wilde and Hess [10], the six most frequently used research methods in descending order are: ADA, CS, PT, QNA, CDA, and FDA. Wilde and Hess [10] define these methods as the six core methods of the German-speaking IS discipline. Our findings confirm the importance of these core methods because these methods have also been the six most frequently used primary research methods in our data sample but in a different order: ADA, QNA, CDA, FDA, CS, and PT. Note that attention should be given when assigning research methods to research paradigms. For instance, in line with Wilde and Hess [10], we assigned research methods, such as the CS, the QNA/QLA, or the LE, to the behavior science research paradigm. However, as our results in section 4.2 reveal, the CS, for instance, has also been frequently used in design-oriented articles for the evaluation of IT artifacts.

In line with Wilde and Hess [10], we also compare our results with the international IS literature. In this context, Wilde and Hess [10] compared their results with the study by Palvia, et al. [20], who examined the use of research methods in the international IS journals between 1993 and 2003. In 2015, Palvia, et al. [21] presented a ten-year update (2014-2013) of this study, in which they analyzed 2487 research articles published in several top-ranked international IS journals. We compare our results with the study by Palvia, et al. [21] because the suggested classification of research methods is similar to the classification of Wilde and Hess [10] and frequency statistics are provided for each research method. Following the argumentation of Wilde and Hess [10], we merged the research methods suggested by Palvia, et al. [21] with our research methods in the following way: The research methods "literature review" and "literature analysis", as suggested by Palvia, et al. [21], were assigned to the argumentative-deductive analysis, which is conceptually similar to these methods. In a similar way, the methods "survey", "secondary data", and "content analysis" were assigned to the quantitative cross-section analysis. The method "interview" was assigned to the qualitative cross-section analysis and the method "field study" was assigned to the case study. The formal-deductive analysis and the simulation were grouped together since Palvia, et al. [21] do not differentiate between these two methods. Moreover, the research method "speculation/commentary" was excluded since we focused in our analysis on articles that provide complete research results (cf. section 3). Figure 8 illustrates the comparison of our study with the study by Palvia, et al. [21]. Following Wilde and Hess [10], we only compared the use of primary research methods.

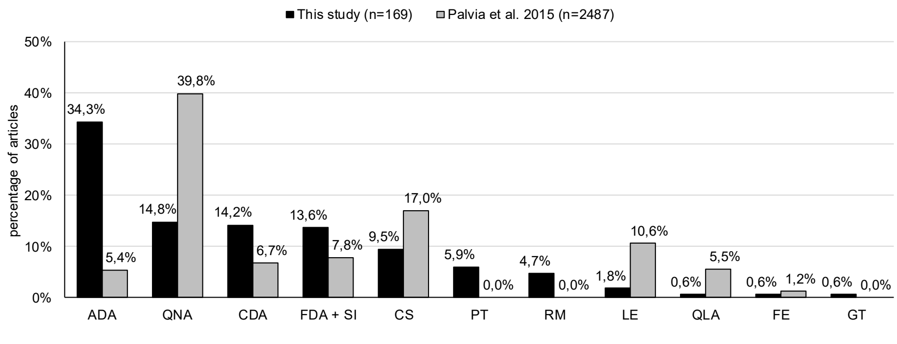


Figure 8. Comparison with Palvia, et al. [21] (only primary methods)

The comparison with the international IS literature reveals that research methods, which are typically associated with design science research paradigm, such as the ADA, the CDA, or the FDA/SI have been quite more often used in our data sample. Interestingly, PT, RM, and GT have not been mentioned in the international IS literature. In contrast, research methods, which are typically associated with the behavior science research paradigm, such as the QDA, the QLA, the CS, or LE, are more often used in the international IS literature than in our data sample. Consequently, our results support the assumption that the design science research paradigm still has a strong position in the German-speaking IS community [10], [11].

5.2 Implications

By illustrating potential shifts in the use of research paradigms and research methods, our results can be used as a basis for future discussion about the positioning and the challenges of the German-speaking IS community. Moreover, our results can also be considered as best practices that support researchers to become more familiar with the profile of the German-speaking IS community. In this way, researchers striving to publish their work in German-speaking IS outlets might better understand how they should position their research and which research methods might best fit to their research projects. For instance, our results show that research articles following the behavior science research paradigm typically employ a quantitative cross-section analysis or a case study. Accordingly, researchers following the behavior science research paradigm can also focus one of these methods or explicitly focus on a less frequently used research method (e.g., grounded theory) in order to generate new insights and further advance the field. Our results also support researchers following the design science research by showing which research methods are typically used in which stage of the design science research process (i.e., build or evaluate) and how these methods can be effectively combined. For instance, our results show researchers following the design science research paradigm that an IT artifact that is built by employing an argumentative-deductive analysis can typically be evaluated through the use of a case study.

5.3 Limitations

Our study has several limitations. First, the German-speaking IS community in this study is only represented by research articles published over the last ten years in the BISE journal. To verify our results, additional journals (e.g., ISeB) and conferences (e.g., WI, MKWI) of the German-speaking IS community have to be examined. Moreover, research articles published in international IS journals (e.g., AIS basket of top journals) and IS conferences (e.g., ICIS, ECIS) should be examined, given the fact that many researchers of the German-speaking IS community publish their work predominantly in the international IS literature.

Second, we focused in our study on the two paradigmatic positions, behavior science research and design science research because these two positions are frequently discussed in the German-speaking IS community [4-7]. However, and as illustrated in section 2, research paradigms can be classified in different ways. Future studies, for instance, can follow the approach of Chen and Hirschheim [12] and examine the use of positivist and interpretative research paradigms in the German-speaking IS literature.

Third, we used the established classification scheme developed by Wilde and Hess [10] to systematically categorize the collected research articles. By doing so, we were able to compare our results directly with the study by Wilde and Hess [10]. While each research method is well described in this classification scheme, it was sometimes not obvious for the research assistants how a particular research article should be categorized. While variations in the results were intensively discussed until a consensus was reached, other researchers might categorize an article in a different way. Nevertheless, we found the classification scheme very robust and we had no reasons for modifications. However, researchers might also apply other classification schemes (e.g., [13]).

Fourth, in our data analysis, we took a more detailed look at the research articles following the design science research paradigm. We made this decision because design science research is considered as the dominant research paradigm in the German-speaking IS community. In a similar way, researchers could also take a more detailed look at research articles following the behavior science paradigm. In this context, research, for instance, could examine average data sample sizes, potential combinations of qualitative and quantitative research methods, or cross-sectional vs. longitudinal studies.

Fifth, we did not consider various environmental factors. For instance, the identified shift in research paradigms (cf. section 4.3) might be a result of the change of the publication language of the BISE journal in 2014 [22]. However, to verify such assumptions, further analyses (e.g., interviews with the journal editors and editorial board members) have to be conducted.

6 Conclusion

In this paper, we summarized which research paradigms and research methods have been employed in the German-speaking IS community, represented by the BISE journal, in the past decade. Doing so enabled us to provide a trend for the last decade on which research methods were most often employed and on how the ongoing positioning discussion of the German-speaking IS community may have led to a change in the diffusion of research paradigms. By examining the use of multiple research methods, we gathered our data in a different way than the study of Wilde and Hess [10]. Accordingly, we were able to depict how many research methods were typically used in one research article and to show in how far the number of methods varied between behavior science and design science articles. Furthermore, we showed which research methods were most often used in combination and which research methods were used to build and to evaluate IT artifacts. With our results, we contribute towards guiding researchers when faced with the decision on how certain IT artifacts could be evaluated or how multiple research methods could be combined in one research approach. Within the light of the IS discipline being an interdisciplinary research field, we, therefore, contribute towards maintaining the paradigmatic and methodological diversity of our community. Moreover, we provide a fruitful and grounded basis for future debates about the positioning of the German-speaking IS community.

References

1. Frank, U.: Herausforderungen der Wirtschaftsinformatik in Zeiten des Wandels. In: Jung, R., Myrach, T. (eds.) Quo vadis Wirtschaftsinformatik, pp. 37-56. Gabler, Wiesbaden (2008)
2. Jarke, M.: Perspectives in the Interplay Between Business and Information Systems Engineering and Computer Science. Business & Information Systems Engineering 1, 70-74 (2009)
3. Winter, R.: What in Fact is Fundamental Research in Business and Information Systems Engineering? Business & Information Systems Engineering 1, 192-199 (2009)
4. Becker, J., Pfeiffer, D.: Beziehungen zwischen behavioristischer und konstruktionsorientierter Forschung in der Wirtschaftsinformatik. In: Zelewski, S., Akca, N. (eds.) Fortschritt in den Wirtschaftswissenschaften: Wissenschaftstheoretische Grundlagen und exemplarische Anwendungen, pp. 1-17. DUV, Wiesbaden (2006)
5. Frank, U., Schauer, C., Wigand, R.T.: Different Paths of Development of Two Information Systems Communities: A Comparative Study Based on Peer Interviews. Communications of the Association for Information Systems 22, 21 (2008)
6. Österle, H., Becker, J., Frank, U., Hess, T., Karagiannis, D., Krcmar, H., Loos, P., Mertens, P., Oberweis, A., Sinz, E.J.: Memorandum on Design-Oriented Information Systems Research. European Journal of Information Systems 20, 7-10 (2011)
7. Becker, J.: Ein Plädoyer für die gestaltungsorientierte Wirtschaftsinformatik. In: Jung, R., Myrach, T. (eds.) Quo vadis Wirtschaftsinformatik?, pp. 3-21. Gabler, Wiesbaden (2008)
8. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28, 75-105 (2004)
9. March, S.T., Smith, G.F.: Design and Natural Science Research on Information Technology. Decision Support Systems 15, 251-266 (1995)
10. Wilde, T., Hess, T.: Forschungsmethoden der Wirtschaftsinformatik. WIRTSCHAFTSINFORMATIK 49, 280-287 (2007)
11. Winter, R.: Design Science Research in Europe. European Journal of Information Systems 17, 470-475 (2008)
12. Chen, W.S., Hirschheim, R.: A Paradigmatic and Methodological Examination of Information Systems Research from 1991 to 2001. Information Systems Journal 14, 197-235 (2004)
13. Palvia, P., Pinjani, P., Sibley, E.H.: A Profile of Information Systems Research Published in Information & Management. Information & Management 44, 1-11 (2007)
14. Steininger, K., Riedl, R., Roithmayr, F., Mertens, P.: Fads and Trends in Business and Information Systems Engineering and Information Systems Research – A Comparative Literature Analysis. Business & Information Systems Engineering 1, 411-428 (2009)
15. Loos, P., Mettler, T., Winter, R., Goeken, M., Frank, U., Winter, A.: Methodological Pluralism in Business and Information Systems Engineering? Business & Information Systems Engineering 5, 453-460 (2013)

16. Benbasat, I., Weber, R.: Research Commentary: Rethinking "Diversity" in Information Systems Research. Information Systems Research 7, 389-399 (1996)
17. Robey, D.: Research Commentary: Diversity in Information Systems Research: Threat, Promise, and Responsibility. Information Systems Research 7, 400-408 (1996)
18. Gregor, S.: The Nature of Theory in Information Systems. MIS Quarterly 30, 611-642 (2006)
19. Hirschheim, R., Klein, H.K.: A Glorious and Not-So-Short History of the Information Systems Field. Journal of the Association for Information Systems 13, 188-235 (2012)
20. Palvia, P., Leary, D., Mao, E., Midha, V., Pinjani, P., Salam, A.: Research Methodologies in MIS: An Update. Communications of the Association for Information Systems 14, 24 (2004)
21. Palvia, P., Kakhki, M.D., Ghoshal, T., Uppala, V., Wang, W.: Methodological and Topic Trends in Information Systems Research: A Meta-Analysis of IS Journals. Communications of the Association for Information Systems 37, (2015)
22. Bichler, M.: Language Change. Business & Information Systems Engineering 6, 317-317 (2014)
23. BISE, Business & Information Systems Engineering (BISE) Journal, http://www.bise-journal.com/ (Accessed: 20.08.2016)
24. Hirschheim, R., Klein, H.K.: Four Paradigms of Information Systems Development. Communications of the ACM 32, 1199-1216 (1989)
25. Weber, R.: The Rhetoric of Positivism Versus Interpretivism. MIS Quarterly 28, (2004)
26. Fitzgerald, B., Howcroft, D.: Competing Dichotomies in IS Research and Possible Strategies for Resolution. In: Proceedings of the 19th International Conference on Information Systems (ICIS), pp. 155-164 (1998)
27. Iivari, J.: A Paradigmatic Analysis of Contemporary Schools of IS Development. European Journal of Information Systems 1, 249-272 (1991)
28. Goles, T., Hirschheim, R.: The Paradigm is Dead, the Paradigm is Dead…Long Live the Paradigm: The Legacy of Burrell and Morgan. Omega 28, 249-268 (2000)
29. Mingers, J.: Combining IS Research Methods: Towards a Pluralist Methodology. Information Systems Research 12, 240-259 (2001)
30. Iivari, J., Hirschheim, R., Klein, H.K.: A Paradigmatic Analysis Contrasting Information Systems Development Approaches and Methodologies. Information Systems Research 9, 164-193 (1998)
31. Orlikowski, W.J., Baroudi, J.J.: Studying Information Technology in Organizations: Research Approaches and Assumptions. Information Systems Research 2, 1-28 (1991)
32. Ebeling, B., Hoyer, S., Bührig, J.: What Are Your Favorite Methods? - An Examination on the Frequency of Research Methods for IS Conferences from 2006 to 2010. In: Proceedings of the 20th European Conference on Information Systems (ECIS), pp. 200 (2012)
33. King, W.R., He, J.: Understanding the Role and Methods of Meta-Analysis in Is Research. Communications of the Association for Information Systems 16, 665-686 (2005)
34. Weber, R.P.: Basic Content Analysis. Sage Publications, Beverly Hills, USA (1990)
35. Baskerville, R., Pries-Heje, J.: Explanatory Design Theory. Business & Information Systems Engineering 2, 271-282 (2010)
36. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly 37, 337-355 (2013)

A Self-Service Supporting Business Intelligence and Big Data Analytics Architecture

Jens Passlick[1], Benedikt Lebek[2] and Michael H. Breitner[1]

[1] Leibniz Universität Hannover, Information Systems Institute, Hannover, Germany
{passlick,breitner}@iwi.uni-hannover.de
[2] bhn Dienstleistungs GmbH & Co. KG, Aerzen, Germany
Lebek.Benedikt@bhn-services.com

Abstract. Self-service Business Intelligence (SSBI) is an emerging topic for many companies. Casual users should be enabled to independently build their own analyses and reports. This accelerates and simplifies the decision-making processes. Although recent studies began to discuss parts of a self-service environment, none of these present a comprehensive architecture. Following a design science research approach, this study proposes a new self-service oriented BI architecture in order to address this gap. Starting from an in-depth literature review, an initial model was developed and improved by qualitative data analysis from interviews with 18 BI and IT specialists form companies across different industries. The proposed architecture model demonstrates the interaction between introduced self-service elements with each other and with traditional BI components. For example, we look at the integration of collaboration rooms and a self-learning knowledge database that aims to be a source for a report recommender.

Keywords: Business Intelligence, Big Data, Architecture, Self-Service, Analytics

1 Introduction

Companies' market capitalization generally consists of enormous amounts of data available to them. However, several companies struggle to use these large amounts of data for analysis or for a decision support as data is often not easily accessible to business users [1]. Business Intelligence (BI) describes the process from collecting data to a fact-based decision support. This decision support is extending from strategic questions into operational environments [2]. This leads to the demand to enable more users to use BI systems. Many companies have to make these decisions in a time-critical environment, which increases the need for a faster technical infrastructure. It is crucial to consider the time a department needs to access the relevant information. Self-service BI (SSBI) provides a solution to these demands. SSBI aims to "empower casual users to perform custom analytics and to derive actionable information from large amounts of multifaceted data without having to involve BI specialists. Power users, on the other hand, can accomplish their tasks with SSBI more easily and quickly than before." [3]

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Passlick, J.; Lebek, B.; Breitner, M. H. (2017): A Self-Service Supporting Business Intelligence and Big Data Analytics Architecture, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1126-1140

Not only the importance of self-service BI rises but also big data analytics is an emerging topic [4]. The increasing volumes of data and the need for advanced analytics means that BI architectures must be adjusted. Many papers discuss parts of a self-service environment but not a whole self-service BI/big data architecture (e.g. [5], [6], [7]). This leads to the following research question:

> *RQ: How is a self-service supporting BI/big data analytics architecture constructed?*

The proposed BI/big data analytics architecture model supports standardized BI reports and new big data analysis, and also enables power users to build their own reports. The research design is described in the next section. After that, the relevant literature is presented. Next, the new architecture and the self-service supporting elements of the collaboration rooms and the knowledge database are explained. Finally, recommendations and implications are given and discussed. Further, the limitations are named and further research is addressed. The paper ends with conclusions.

2 Research Design and Methods

In order to ensure methodological rigor, this study utilizes design science research as the underlying methodology as it is well suited for the development of an architecture. Mainly we were guided by the Design Science Research Model (DSRM) proposed by [8]. Figure 1 shows the phases and the steps that were carried out. Using a literature review based on Webster and Watson, relevant BI and big data architecture models were discussed and a research gap was identified [9]. In the next step ("Objectives definition") SSBI literature was analyzed and demands from practice were included. With these insights a conceptual model was developed. Open semi-structured interviews helped to improve the model in the "design & development" phase. This research method makes a free discussion about the problems and requirements of SSBI possible. Eighteen experts from different industries were interviewed (see table 1). Each expert had at least two years of experience with BI and on average, they had ten years. The interviews lasted on average one hour. The interviews were transcribed and analyzed by categorizing the main statements. Mayring's method makes qualitative statements comparable by analyzing the frequency in which they were mentioned [10]. The improvements were incorporated and the changed model was shown to the experts again. The new improvements were implemented in the next step.

Table 1. Interviewed experts

Job Group	*Expert Number*
Business consultant	1-3
SAP consultant	4-8
BI application developer	9-13
IT manager	14-18

The demonstration and evaluation phase of the original Peffers et al. model was summarized with an applicability check [11]. A focus group consisting of eleven researchers and a group consisting of twelve practitioners discussed the model with regard to whether it adds value for research and practice and whether it can specifically help in the implementation of SSBI.

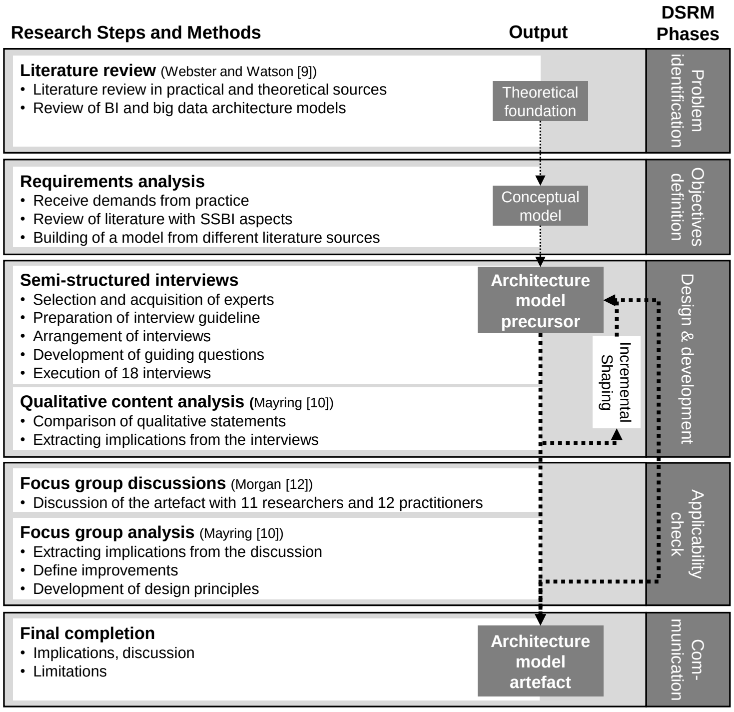


Figure 1. Research design based on [8]

3 Proposed Self-Service Supporting Architecture

3.1 Status Quo and Problem Identification

To identify the status quo of the SSBI research, a literature search was done in the AISeL, ScienceDirect, IEEEXplore, ACM and Emerald database. It was extended to include practitioner resources. Whitepapers by the BeyeNETWORK, The Data Warehousing Institute (TDWI), and Gartner were analyzed. The search keywords we used contained: “Self-Service” in combination with “BI”, “Business Intelligence”, “Big Data”, “Architecture” and “Analytics.” The publication dates ranged from 2005 to the present. The search resulted in 1,258 potentially relevant articles. They were reviewed by title and unsuitable papers were eliminated. If the title did not make a clear decision possible, the abstract, the introduction, and the conclusion were consulted. After that a forward and backward search in the most relevant papers was conducted. This included non-academic literature like whitepapers. Forty articles were deemed highly relevant for the development of the model. The literature review identified eight different BI or big data architecture models. Phillips-Wren et al. propose a big data analytics architecture model based on different other models [4]. The authors analyze existing BI/big data literature and describe a new user group they call data scientists. In the field of data processing infrastructures, Phillips-Wren et al. focus on the use of Hadoop clusters as a solution for big data use cases. Another model proposes a service-orientation character for a BI architecture [13]. They developed a BI architecture model that shows how this service character is implemented and which elements are necessary. Their model does not consider big data analytics use cases in particular. Another model is provided by [46]. In their work they focus on a mapping layer and a semantic layer which should be between the users and a data warehouse. A paper by Imhoff describes the different tracks for data processing in a big data environment [14]. It is a similar idea to the concept of a lambda architecture [15]. None of the previously discussed models make any statements about SSBI. The models by Watson and Eckerson provide some ideas for an implementation of SSBI [16-17]. Watson improved on Eckerson's model. The two models illustrate the difference between top-down and bottom-up BI. Top-down BI describes a BI environment that is very predefined and fixed whereas Bottom-up BI is an open environment that is not predefined [18]. Both models only differentiate between two user groups. [19] developed a model with a focus on SSBI. Their model describes different data processing methods, has a semantic layer, and covers big data analytics use cases. But they do not deal with different user groups. Another concept is to support SSBI with a business level ontology [20]. This is supposed to make the data model more comprehensible for the end user. [21] also propose a semantic layer to realize a unified business view of the data.

3.2 Requirements: Existing SSBI Aspects in Literature

In the second phase of the research design the objectives have to be defined. This is done by reviewing additional literature describing certain aspects of the implementation of SSBI. They can be separated into five groups: Special SSBI governance aspects and guidelines, concepts for an individual BI usage, social media elements in a BI environment, collaboration concepts and concepts for a knowledge database. This is summarized in table 2. Papers with special SSBI governance aspects and guidelines deal with changes in BI/big data analytics governance strategies [22], different ETL ("extract", "transform" and "load") processes [23], the need of special tools [24] or SSBI guidelines [25], [21], [26]. The individual BI usage group includes papers which describe concepts for an individual use of the BI environment. The idea of the integration of social media elements into a BI environment is to support the usage and the collaboration of BI users. Collaborative BI comprises the cooperation in the creation of reports or queries. In this context, it means human cooperation and not the grouping of systems. It is stressed that collaborative BI is not simply an element that has to be implemented into a BI architecture in terms of a technical platform; it also has to begin in the minds of employees [27]. The last group of papers considered deals with a knowledge database. The idea behind it is that the construction of every analysis and report is saved in an additional database. This includes the history of the conducted analyses and the order of their execution. Through that, forecasts of analysis paths should be possible.

Table 2. Overview SSBI literature

SSBI aspects	*Description*	*Sources*
Governance and guidelines	Changes in governance and guidelines for the realization of SSBI	[21-26]
Individual BI usage	Concepts which support an individual BI usage	[5-7], [28-30]
Social media elements	Social media elements in a BI environment	[31-32]
Collaboration	Cooperation in the creation of reports or queries	[5], [25], [27], [33-35]
Knowledge database	Database which saves construction and usage of reports and analyses; also examination of analysis paths	[36-41]

A combination of these elements with a comprehensive BI/big data analytics architecture is still missing. In the following, the focus will be on the architecture itself, the implementation of collaboration rooms, and a self-learning knowledge database. The collaboration rooms can then be connected with existing enterprise social media systems. After developing a first model with the findings from literature the model was improved through expert interviews. The following table 3 shows some of the major changes caused by the expert interviews.

Table 3. Improvements through expert interviews

Model layer	*Description*	*Sources*
Preparation	Multiple data access methods added; added direct access without using a storage system	Experts 1, 5, 7, 11, 13
Storage and analysis infrastructure	Generalization of the storage and analysis infrastructure into three tiers	Experts 1, 3, 6, 15-16
Presentation	Enterprise social networks added, skills added	Experts 4, 11
Knowledge database	Feedback loop added, development of the different use cases of the knowledge database	Experts 2, 4, 12, 15, 16
Governance	Order of the governance aspects according to by the experts mentioned importance	All experts had influence

3.3 Model Overview

In the following, the final model developed with the help of expert interviews is explained. Inspired by existing BI/big data analytics architecture models, the aim is to describe the whole process from the data sources through to the presentation of information. Big data is defined as “a phenomenon characterized by an ongoing increase in volume, variety, velocity, and veracity of data that requires advanced techniques and technologies to capture, store, distribute, manage, and analyze these data.” [42] This is the reason for the need of an advanced technical infrastructure. The changed technical infrastructure leads to a more complex data access for users which effects the possibilities of SSBI and the need to discuss the entire BI process from the source systems to the presentation of the data.

On the left side of the model are the data sources. The data sources are separated into internal and external sources. The data origin shown in Figure 2 are examples of those sources. The next step in data processing is the preparation of the data. Three different ways of accessing data exist. The first one is a direct access tunnel for analysis, where a special integration or caching of the data is not necessary. Second, direct access for real-time analysis is shown. The third method is a classic ETL process. But this process is extended by the possibility of performing an EL(T) process [43]. EL(T) stands for "extract", "load" and an optional "transform" process. This takes into account that in some big data analysis there can be a need for raw data that is not transformed. Different data access methods have to be taken into account for realizing SSBI. This is especially important for data scientists, who need access to raw data. In the proposed model, the storage and analysis infrastructure layer consists of two main and one optional tier. An element for data integration is necessary in every BI or big data environment. The job can be done with a classic data warehouse, but other technologies can take on this job, such as in-memory databases or Hadoop clusters. The other tier is the "big data refinery." This element ensures the necessary infrastructure for big data analysis and includes "experimental platforms." These platforms are essential for the data scientist user group. They need possibilities for experiments where data from different sources can be staged, merged, and analyzed [44]. The last tier consists of optional elements that could be necessary for a real-time BI realization, such as data caches [45]. To simplify access to data across multiple systems, there is the semantic layer which also includes the mapping layer described in [46]. It realizes a unified access to the different storage systems and an easier access to the data for users with low technical skills. A possible embodiment of the semantic layer could be a service oriented architecture. A service oriented BI architecture is described in the work by Pospiech and Felden [13].

The presentation of the data is separated into three portals. This separation is done according to the skill and the need of the BI user. In the dashboards, the users are consumers of predefined reports and they have a low degree of freedom [16-17]. Dashboards are mainly used by casual users. On the other side is the group of data scientists. In their data laboratory they have a high degree of freedom, as well as the access rights and tools to completely build their own analysis and reports. As mentioned above, they need platforms for experiments with new analyses because they are dealing with large and unstructured data sets. Between those two platforms the analytics portal is located. This is the main platform for SSBI applications. Reports are predefined but users can adjust the reports with restrictions. In general, the experts agreed with this representation. Some experts had a slight different user group definition in their own company like Watson also distinguishes between five user groups [44]. These user group definitions can therefore be adapted to the individual needs of the respective company. This is expressed in the following quote: *"Sure, there might be sub-groups, especially within the group of the power users and in the data scientists. But I think with three groups it is quite concise. Those are the right groups in the model. It is also* *meaningful to distinguish the groups by the user skill."* (Consultant, mid-sized consulting firm - interview conducted in German)

In addition to the definition of the different user groups, one expert added that the interaction between the portals plays an important role in supporting SSBI. *"The transition must be very fluent. The dashboard must be easy to use and allow a simple jump into details. So you need to have a drill down functionality. The continuity is important and just the same the usability. One must like to use the tool or the portal, because it is easy to use."* (Head of a quality management department, mid-sized industrial company - interview conducted in German)

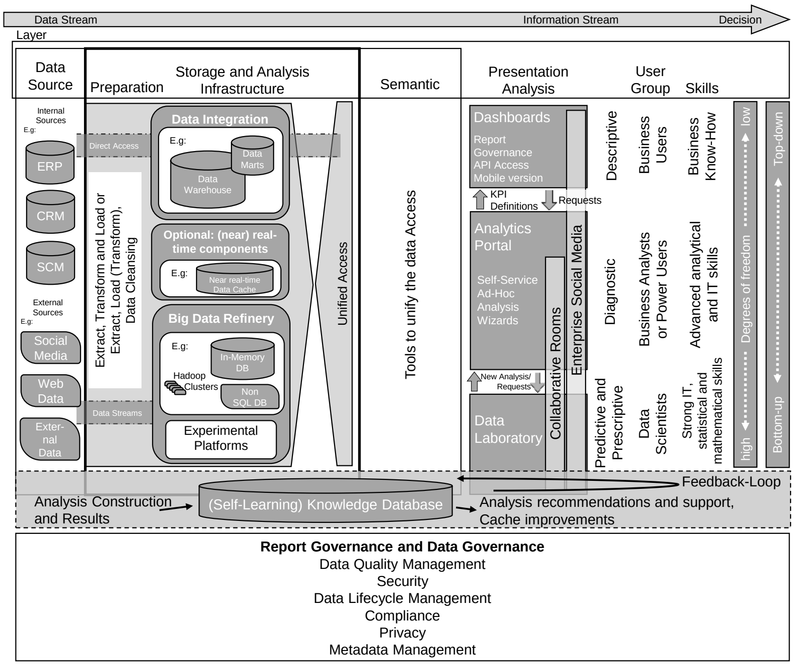


Figure 2. Proposed architecture model

To support the interaction the developed model is connected to an enterprise social network of the respective company. In that way the collaboration rooms can be merged with the enterprise social network and the exchange between the user groups can be encouraged. Below the model different aspects of a report and data governance are mentioned. They are ordered according to a ranking by the experts.

After giving a rough overview of the model the two main elements for the support of SSBI are described.

3.4 Collaboration Rooms

The "collaboration room" architectural component is a platform where a direct cooperation from users of the analytics portal and the data laboratory is possible. Users of the same portal can cooperate while working on the same platform. Also, users of the analytics platform can give feedback for analyses performed by data scientists. Business analysts can also ask for special sub-parts of their analysis to be built by data scientists. It is important for the process that the collaboration history is saved. Today most collaboration communication is done by email. The problem is that only the people involved have access to the origin story of a decision-making process. A collaboration platform can replace email communication. [25]

Figure 3 shows proposed classes of a collaboration room environment. It represents the different user groups and the related platforms. Business users and analysts can create requests for a new report or analysis. Business analysts can also ask for help with the construction of a report. The collaboration can take place inside a user group or business analysts can make requests to data scientists. These requests are connected to one or more reports. Every report belongs to a workspace. This is the main room where the collaboration can take place. Inside a workspace it is possible to create several communication rooms. One-on-one and group discussions are possible. The workspaces in conjunction with the communication rooms provide the opportunity for discussing reports, creating different report versions, and conducting experiments. All these elements support the collaboration between the different user groups of the BI/big data analytics architecture.

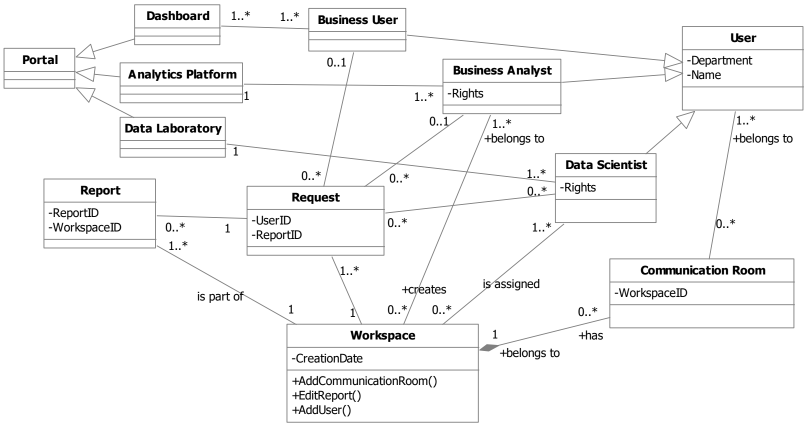


Figure 3. Collaboration environment conceptual class chart

3.5 Self-learning Knowledge Database

The knowledge database saves all performed queries except special experiments in the data laboratory. This includes the results of the queries as well as the queries themselves. It is useful to keep the queries for later use because they can have enormous value for later analyses. An historic analysis database creates the possibility for the replication of an analysis, which makes the building of new complex analysis easier. Here a service orientation shows its advantages because it is possible to easily see which components and services were used by different analyses. There is also added value generated by the possibility of showing related analyses [35]. This helps a business analyst build a new analysis or find further queries that were created in the past or by another user. [36], [39]

After conducting the expert interviews, several reasons for the introduction of a knowledge database were identified. The main reason is to improve SSBI with recommendations for similar analyses or by supporting the developing process of analyses [35], [38]. It can help to improve dashboards because with the database, it is known how often a report was accessed. Another important point is that the knowledge database helps to fill a cache in advance. This is made possible by the self-learning mechanism, which allows predictions. If we know which analyses are accessed frequently, the results can be computed in advance and saved into the cache. Then, fewer calculations have to be computed because the results are already in the cache, which decreases the response time. The prediction of queries can also be done by using Markov models [41].

Figure 4 shows a class chart of the proposed knowledge database. It represents the three potential use cases: Help while building new reports or queries, recommender for further analyses that might be interesting for the user, and intelligent filling of caches. A user builds or calls an analysis. This call is written into the knowledge or meta database just like the analysis path. The analysis path consists of the order and the connected queries a user calls in a session [40]. With the learning engine, all the data from the meta database is analyzed. Intelligent algorithms look for relations inside the queries and between the analysis paths. Different learning engines with different algorithms are possible.

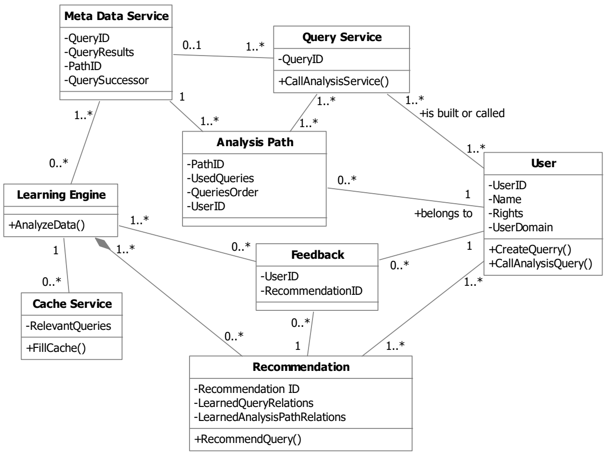


Figure 4. Knowledge database conceptual class chart

A problem could be that big amounts of data are necessary for meaningful results. This was already discussed with the experts: *"What you really need is: First you need a lot of different executions on top of your system. So it won't work in a single enterprise because you won't have enough data for your analytics of the analysis templates or mechanism that work and you need very good feedback functions. So you need to look in the usage data. So what is used and what is successful. […] I think that it will be very hard to build it on premise. It is something that works in pretty large companies because* *otherwise there is not enough usage for this."* (Vice President Platform, cloud BI provider) The results of the learning engine is then used by the recommender engine and the cache service. For the learning process it is essential to have a feedback loop. This means that the user can evaluate the results of recommendation. This feedback is then used by the learning engine for the improvement of the recommendation processes.

4 Implications, Recommendations and Discussion

The aim of academic literature is not only to focus on theory, but also to provide relevance for practitioners in order to prevent research from becoming an end unto itself [11]. For this reason, an applicability check was done after the final model was developed. In two focus groups consisting of practitioners and researchers, it was discussed whether the model can help realizing SSBI. In general, it was stated that the architecture model is helpful because it reduces complexity and gives companies a point of orientation. It was further remarked that for an application in practice, it must be defined further which use cases are relevant for SSBI. The focus groups also discussed potential main user groups in a SSBI environment. The discussion participants thought that it might be a user group that has ad-hoc questions that are not regular. They stated the need of a semantic layer and discussed that a service orientation as described in the model [13] is useful, but it requires a high degree of standardization in processes for the acquisition of information. This is a big problem in companies because these processes are mostly unknown. In the company of the focus group members the aim is to use '*Business Objects Universes'* for the realization of a semantic layer instead of realizing a complete service oriented architecture. In further projects '*SAP Business Objects Design Studio'* will be used for the creation of dashboards and '*Analysis for Office'*, an Add-In for *'Microsoft Excel'*, should be used for the analytics portal of the architecture. A data laboratory is not planned at the moment. The knowledge database can help new users in particular because they can get an idea of what information is available. This is supported by the statement of one expert: *"A typical use case for our big customers is that if you are a user and you create a new report then you have 99% chance that somebody else has already done this report. Exact this report! So that is the simplest thing. You can just search the report, look at the structure, look at the dimensions or whatever components the user is working with and start with what is already there. The second thing is some kind of recommendation. In our case it can be driven by what people will be doing."* (Vice President Platform, cloud BI provider)

Another point is the meaningfulness of the collaboration rooms. A different expert describes the value that is generated through a well-organized collaboration, but notes that there are still good implementations missing. *"I think this is valuable and useful because I think this should be the way into the future. Get out of the habit of each person making his or her own report, but that you can also reuse more of the reports. […] However, in reality it is not so simple to find platforms that make the realization possible. I have not yet seen and experienced properly implemented collaborative rooms in practice."* (Consultant for SAP BI, consulting firm - interview conducted in German) Table 4 summarizes the findings of this research in design principles. Besides these design principles the main output of this research is the architecture model which is presented in Figure 2. It shows the interaction between the elements and their position in the BI/big data analytics process.

Table 4. Design principles

Architecture element	*Design principle*
Data access	The data access via different data sources should be simplified by a unified access. This paper proposes a service orientation for the realization.
Semantic layer	To achieve a unified access to the data there should be a semantic layer which connects the different data sources. This could be in a service oriented but the applicability check showed that other realizations are possible, too.
User groups	To address the individual needs of the BI users, a definition of different user groups is necessary. This paper proposes three different groups but point out that this has to be adjusted according to the structures of the respective company.
Different portals	To address the different needs of the user groups, different portals are suggested.
Collaboration	Collaboration opportunities should be considered in a BI/big data analytics architecture. Enterprise social media can support the collaboration in a BI/big data environment.
Knowledge database	A knowledge database should be used in conjunction with a service oriented architecture to assist new users, for an intelligent cache usage and to help users with building new reports.

In the following section, the results of this research are described and compared with existing work. In terms of a semantic layer, as proposed in the literature [28], [30], the developed model stays universal but sees advantages in a service-oriented approach [13]. This is a concrete solution and it is assumed that this service orientation can be handled well in the knowledge database. Elements like a service repository are seen as being useful in supporting SSBI. Some experts criticized the fact that a service orientation would require a lot of effort in the beginning to standardize all the processes and services. The focus group decided that this might be a general problem of SSBI. The right balance must exist between standardization and flexibility.

In a big data analytics architecture, a new storage and analysis infrastructure is necessary. This paper connects the idea of many big data contributions [4], [17-16], [19] and assigns the new technologies to three tiers, similar to other proposals [14-15]. Especially the big data refinery in conjunction with experimental platforms are important for the independent work of the data scientists. The presented user groups are similar to the definitions of other research [4], [6]. The expert interviews showed that these definitions can be found in practice, but the probability of deviations in practice is high. Therefore, it is important to know the user groups of the BI architecture in order to correctly address the individual needs of each user group in an SSBI context. As mentioned by one of the experts, there is a need of fluent transitions between the portals.

The knowledge database can also support SSBI. It can contribute to an intelligent filling of analysis caches [41] and can recommend further analysis for users, which especially helps new and unexperienced users [36-39]. Analysis knowledge can be preserved with the use of the knowledge database. Nevertheless, the self-learning function is only realizable if enough data is available. This restricts the use of the self-learning function to large companies or to the use in a cloud environment. Research is moving towards presenting a class chart to give a better idea of how a self-learning knowledge database can be built. This enables storing implicit knowledge of BI users which facilitates an increased value for companies. This is supported by the results by Kretzer et al. [39] who find out that the ease of use of a BI platform is higher with a recommendation system. But they do not consider historical data and therefore they not have included learning loop. The developed collaboration rooms are based on the paper by Berthold et al. [25]. A more concrete implementation possibility is presented and the different user roles are shown. [33] describe another approach with the reformulation of queries in a peer-to-peer network. The collaboration rooms can be seen in connection with enterprise social media elements [31-32]. The expert interviews and the focus group discussion confirmed that the value of collaboration in a BI/big data analytics context will increase and collaboration rooms are a solution for that. It is suggested that collaboration rooms are included into companies' BI/big data analytics strategies as presented in Figure 3.

The whole architecture helps companies define their expectations of a BI/big data analytics architecture. When comparing an existing architecture with the proposed model, weak points and improvement potential can be shown.

5 Limitations and Future Work

A rigorous literature review was conducted. Nevertheless, this method has limits. The search was only done with keywords in English. Publications in other languages could not be considered unless they could be found by means of a forward or backward search. Eighteen interviewees were asked for critique and improvements. To obtain more objective opinions, a larger number of interviews would be useful. This could reduce the likelihood that important aspects are forgotten or misrepresented. The background of the experts is rather homogeneous. Most experts are BI consultants or BI developers who are good at discussing the overall architecture. The opinions of business users are still missing. Further research could use these opinions to improve the design of the presentation layer. Especially in relation to SSBI, their view might still be a significant enhancement. Furthermore, different business sectors were not considered. Thus, no statements about the adaption of the architecture to specific branches is possible. Further research can be done by asking how the architecture must be adjusted according to a special domain or how different sizes affect the architecture. It is obvious that the architecture has to be adjusted individually for every company. It was mentioned by the experts, for example, that the realization of a self-learning knowledge database is highly dependent on the amount of potential input data. If a company is not big enough to provide a sufficiently large amount of reports, inquiries, and analyses, there might not be enough input data. The recommendation function is then very limited because the learning algorithm does not get enough data. In such a case, only a simple knowledge database without self-learning algorithm could be realized. Statements about the amount of required data for a good working self-learning BI recommendation algorithm are not possible. Further research is needed with respect to that area. A self-learning knowledge database prototype should be developed. This could enable further discussions on this issue. The alternative would be a cloud-based knowledge database. By analyzing the reporting and analysis paths of several companies, a cloud implementation could deliver meaningful recommendations. But for that to happen, many architecture components would have to be moved to the cloud. Another question is which algorithms can be used to get useful results out of this analysis. The inclusion of the actual decision into the BI process is an outstanding research question. In this context, it could not be explored to what extent the actual decision can be included in the knowledge database.

6 Conclusions

The developed model shows significant progress in relation to other proposals [4], [16]. It is extended especially with regard to SSBI. The ideas result from both practical and academic literature and in particular from interviews with experts. A focus group discussion was used to check the practicability of the model. The new model represents a universal BI / big data analytics reference model. It can be seen as a guideline for companies, who can evaluate their existing architecture with the aim of improving their SSBI or big data analytics capabilities. It takes different user groups and their different demands into account in a BI/big data analytics architecture. Collaboration rooms and a (self-learning) knowledge database are presented as additional supporting elements. Discussions with practitioners have shown that these elements have great potential to support SSBI because they make implicit knowledge of BI users usable. In further research the applicability should be reviewed by various companies.

References

1. Kosambia, S.: Business Intelligence The Self-Service Way. DM Review 18, 7, 20–22 (2008)
2. Böhringer, M., Gluchowski, P., Kurze, C., and Schieder, C. A: Business Intelligence Perspective on the Future Internet. AMCIS 2010 Proceedings, (2010)
3. Alpar, P. and Schulz, M.: Self-Service Business Intelligence. Business & Information Systems Engineering 58, 2, 151–155 (2016)
4. Phillips-Wren, G., Iyer, L., Kulkarni, U., and Ariyachandra, T.: Business Analytics in the Context of Big Data: A Roadmap for Research. Communications of the Association for Information Systems 37 (2015)
5. Abelló, A., Darmont, J., Etcheverry, L., et al.: Fusion Cubes: Towards Self-Service Business Intelligence. International Journal of Data Warehousing and Mining 9, 2, 66–88 (2013)
6. Corral, K., Schymik, G., Schuff, D., and Louis, R.S.: Enabling Self-Service BI through a Dimensional Model Management Warehouse. Proceedings of the 21st Americas Conference on Information Systems (2015)

7. Varga, J., Romero, O., Pedersen, T.B., and Thomsen, C.: Towards next generation BI systems: The analytical metadata challenge. In Data Warehousing and Knowledge Discovery. Springer International Publishing, 89–101 (2014)
8. Peffers, K., Tuunanen, T., Rothenberger, M., and Chatterjee, S.: A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24, 3, 45–77 (2007)
9. Webster, J. and Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly 26, 2 (2002)
10. Mayring, P.: Einführung in die qualitative Sozialforschung. Beltz Verlag (2002)
11. Rosemann, M. and Vessey, I.: Toward Improving the Relevance of Information Systems Research to Practice: The Role of Applicability Checks. MIS Quarterly 32, 1, 1–22 (2008)
12. Morgan, D.L.: Successful Focus Groups Advancing the State of the Art. SAGE Publications, Inc (1993)
13. Pospiech, M. and Felden, C.: Service-Oriented Business Intelligence Reference Architecture in Face of Advanced BI Concepts. 19th Americas Conference on Information Systems, AMCIS 2013 (2013)
14. Imhoff, C.: Ensuring a Sustainable Architecture for Data Analytics. Intelligent Solutions, Inc., (2015)
15. Marz, N. and Warren, J.: Big Data: Principles and best practices of scalable realtime data systems. Manning (2015)
16. Watson, H.J.: Tutorial: Big Data Analytics: Concepts, Technologies, and Applications. Communications of the Association for Information Systems 34, 24 (2014)
17. Eckerson, W.: Big Data Analytics: Profiling the Use of Analytical Platforms in User Organizations. BeyeNetwork, (2011)
18. Eckerson, W.: ANALYTIC ARCHITECTURES: Approaches to Supporting Analytics Users and Workloads. BeyeNetwork, (2011)
19. Parenteau, J., Sallam, R.L., Howson, C., Tapadinhas, J., Oestreich, T.W., and Schlegel, K.: Technology Insight for Modern Business Intelligence and Analytics Platforms. Gartner (2015)
20. Spahn, M., Kleb, J., Grimm, S., and Scheidl, S.: Supporting business intelligence by providing ontology-based end-user information self-service. Proceedings of the First international Workshop on ontology-Supported Business intelligence, ACM (2008)
21. Schlesinger, P.A. and Rahman, N.: Self-Service Business Intelligence Resulting in Disruptive Technology. Journal of Computer Information Systems 56, 1, 11–21 (2015)
22. Meyers, C.: How Data Management and Governance Can Enable Successful Self-Service BI. Business Intelligence Journal 19, 4, 23–27 (2014)
23. Howson, C.: Embrace Self-Service Data Preparation Tools for Agility, but Govern to Avoid Data Chaos. GARTNER, (2015)
24. Powell, R.: The Convergence of BI, Big Data, Analytics and Mobile Technologies: A Spotlight Q&A with Michael Corcoran and Jake Freivald of Information Builders. BeyeNetwork, http://www.b-eye-network.com/view/17066 (Accessed: 17.08.2016)
25. Berthold, H., Rösch, P., Zöller, S. et al.: An architecture for ad-hoc and collaborative business intelligence. Proceedings of the 1st International Workshop on Data Semantics - DataSem '10 (2010)
26. Smuts, M., Scholtz, B., and Calitz, A.: Design Guidelines for Business Intelligence Tools for Novice Users. Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists (2015)
27. Imhoff, C. and White, C.: Collaborative BI: Theory Becomes Reality. Business Intelligence Journal 18, 2, 40–46 (2013)

28. Demirkan, H. and Delen, D.: Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. Decision Support Systems 55, 1, 412–421 (2013)
29. Kobielus, J., Karel, R., Evelson, B., and Coit, C.: Mighty mashups: do-it-yourself business intelligence for the new economy. Forrester Research (2009)
30. Zorrilla, M. and García-Saiz, D.: A service oriented architecture to provide data mining services for non-expert data miners. Decision Support Systems 55, 1, 399–411 (2013)
31. Alpar, P., Engler, T.H., and Schulz, M.: Influence of social software features on the reuse of Business Intelligence reports. Information Processing and Management 51, 3, 235–251 (2015)
32. Böhringer, M., Gluchowski, P., Kurze, C., and Schieder, C.: On the Role of Social Software Techniques for the Design of Self- Organising Enterprise Reporting Portals. Proceedings of the ITI 2009 31st International Conference on Information Technology Interfaces, 153–158 (2009)
33. Golfarelli, M., Mandreoli, F., Penzo, W., Rizzi, S., and Turricchia, E.: OLAP query reformulation in peer-to-peer data warehousing. Information Systems 37, 5, 393–411 (2012)
34. Kaufmann, J. and Chamoni, P.: Structuring Collaborative Business Intelligence: A Literature Review. 47th Hawaii International Conference on System Sciences, 3738–3747 (2014)
35. Kretzer, M. et al.: Design Principles for Diffusion of Reports and Innovative Use of Business Intelligence Platforms. Wirtschaftsinformatik Proceedings, 675–690 (2015)
36. Aligon, J., Gallinucci, E., Golfarelli, M., Marcel, P., and Rizzi, S.: A collaborative filtering approach for recommending OLAP sessions. Decision Support Systems 69, 20–30 (2015)
37. Giacometti, A., Marcel, P., and Negre, E.: A Framework for Recommending OLAP Queries. Proceedings of the ACM 11th international workshop on Data warehousing and OLAP, 73–80 (2008)
38. Jerbi, H., Ravat, F., Teste, O., and Zurfluh, G.: Applying Recommendation Technology in OLAP Systems. Enterprise Information Systems 24, 220–233 (2009)
39. Kretzer, M. et al.: Designing a Report Recommendation Assistant: A First Design Cycle. International Conference on Design Science Research in Information Systems, 9073, 87–103 (2015)
40. Mertens, M. and Krahn, T.: Knowledge Based Business Intelligence for Business User Information Self-Service. In Collaboration and the Semantic Web, 271–296 (2012)
41. Sapia, C.: PROMISE: Predicting Query Behavior to Enable Predictive Caching Strategies for OLAP Systems. Proceedings of the Second International Conference on Data warehousing and Knowledge Discovery (DAWAK 2000), 224–233 (2000)
42. Ebner, K., Buhnen, T., and Urbach, N.: Think Big with Big Data: Identifying Suitable Big Data Strategies in Corporate Environments. 47th Hawaii International Conference on System Sciences, 3748–3757 (2014)
43. Dmitriyev, V., Mahmoud, T., and Marín-Ortega, P.M. SOA enabled ELTA: approach in designing business intelligence solutions in Era of Big Data. International Journal of Information Systems and Project Management 3, 3, 49–63 (2015)
44. Watson, H.J.: Data Lakes, Data Labs, and Sandboxes. Business Intelligence Journal 20, 1, 4–7 (2015)
45. YiChuan, S. and Yao, X.: Research of Real-time Data Warehouse Storage Strategy Based on Multi-level Caches. Physics Procedia 25, 2315–2321 (2012)
46. Villegas-García, M.A. et al.: How Business Analytics Should Work. In: Advanced Business Analytics. pp. 93–108 Springer International Publishing (2014)

Towards a Capability Model for Big Data Analytics

Christian Dremel[1], Sven Overhage[2], Sebastian Schlauderer[2], and Jochen Wulf[1]

[1] University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland
{christian.dremel,jochen.wulf}@unisg.ch
[2] University of Bamberg, Chair of Industrial Information Systems, Bamberg, Germany
{sven.overhage, sebastian.schlauderer}@uni-bamberg.de

Abstract. Big data analytics is becoming a veritable source of competitive advantage as it helps companies to better understand their business environment and to create or improve their products and services accordingly. However, big data analytics also poses challenges to organizations with respect to establishing the required capabilities. Building upon a design science research approach and the Work System Theory as a kernel theory, we identify several capabilities necessary to leverage the potential of big data analytics. To achieve this goal, we conducted 16 interviews with experts from an IT-strategy consulting firm. We furthermore organize the identified capabilities into a coherent model. The resulting capability model consists of eight capability groups that contain 34 capabilities. It provides a basis to systematically develop the necessary capabilities for the adoption und strategic usage of big data analytics.

Keywords: big data analytics, capability model, work system theory, design science

1 Introduction

The accumulating evidences of potential benefits provide a legitimation to consider big data analytics (BDA) a sustainable phenomenon rather than a buzzword [1], [2]. BDA opens up new business opportunities, such as using “real-time information from sensors, radio frequency identification and other identifying devices to understand business environments at a more granular level, to create new products and services, and to respond to changes in usage patterns” [2, p. 22]. The implementation of BDA poses challenges especially regarding the development of appropriate organizational competencies [3], [4], because the “expanding sea of data […] is either too voluminous or too unstructured to be managed and analyzed through traditional means” [2, p. 22].

Those challenges originate from the vast amount of data that comes in both structured and unstructured forms and from various sources such as the Web, social media, or the Internet of Things [5], [6]. This leads to specific and novel implications for organizations on a procedural (e.g., new forms of decision-making), organizational (e.g., new employee competencies and new structures) and technological (e.g., new platforms and tools) level. Accordingly, the adoption of BDA requires organizational transformations as well as the development of specific analytical and technological capabilities

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Dremel, C.; Overhage, S.; Schlauderer, S.; Wulf, J. (2017): Towards a Capability Model for Big Data Analytics, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1141-1155

(e.g., the effective deployment within the current IT landscape) [1], [7], [8]. For instance, the capability development encompasses the acquisition of sufficient knowledge of how to extract business value from big data and the application and management of the underlying technologies [2], [9]. Companies have a tough time transforming towards a data-driven company and recognizing their data not as a side-product but as a source of competitive advantage. The development and adoption of capabilities represents a first step in following this path. To facilitate this journey, maturity or capability models provide guidance for companies to assess their current situation regarding the capabilities required for a task (i.e., BDA) [10]. However, to the best of our knowledge no capability model exists in IS research, which helps companies to develop and manage big data analytics competencies [11]. In order to address this apparent research gap, we pose the following research question: "*Which capabilities are required to build a big data analytics competence?*" A capability model for BDA will help to assess the current state of a company and to identify necessary initiatives to build required capabilities. To develop such a model in a rigorous scientific process, we follow the design science paradigm. Building upon the Work System Theory as kernel theory and the results of 16 expert interviews, we present 34 capabilities and summarize them in an initial version of our model that will be developed further in future iterations.

Next, we describe the theoretical background and related work. In section 3, we present our research approach. Thereafter, we discuss the interview results and describe the resulting capability model. We conclude by discussing future research avenues.

2 Background and Related Work

To set the background for our research endeavor, we first reflect BDA from a capability perspective, before we discuss related work that is potentially relevant for our research.

2.1 Big Data Analytics as a Set of Capabilities

BDA does not only stand for technological but also for organizational characteristics (e.g., data-driven culture) and their potential to be a source of competitive advantage [4], [12]. From a technological standpoint, BDA as a competence can be characterized by processing requirements with respect to data volume, velocity, variety, and veracity [23]. Big data itself is regarded as the result of all these dimensions [13]. However, conceptualizing BDA purely from a technological perspective does not sufficiently characterize the potential role of BDA in companies [14]. Indeed, BDA requires a data-driven culture, new analytical methods, competencies and capabilities [15]. It "forces us to look beyond the tried-and-true methods that are prevalent" [16, p. 44] today and is characterized by the "belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy" [17, p. 663]. Data itself is useless if no process of sense-making takes place and the gained insights are not used to make data-based decisions [18], [19]. Successfully translating big data-drawn insights into convincing arguments for managers changes the dynamics of managerial decision-making [5]. In this context, the development of specific BDA capabilities can be seen as a major challenge for companies, especially for those digitizing their business models [8].

BDA has a transformative effect on the organization (e.g., substitution of inefficient business models) or competitive landscapes (e.g., entrance of new players, shift of power), which in turn leads to better organizational performance [1], [3], [8], [13]. BDA uncovers previously unknown patterns, correlations and information and can be seen as a solution for gaining better insights from diverse und previously unexploited data sources (e.g., social media, wearables, RFID) [20]. Organizations nowadays realize that analyzing big data can help to stay competitive in terms of profits, speed, efficiency and customer-orientated service through timely and profound decisions [12]. BDA hence poses a potential competitive advantage [1], [12], [13]. However, before this advantage can be achieved, organizations need to analyze how they can use BDA to improve their business and how they can establish necessary organizational capabilities [19], [20]. Organizational capabilities, in general, comprise skills, abilities and expertise of an organization and they are idiosyncratic and inimitable [21]. Representing a source of organizational value and leading to competitive advantages, they are connected tightly to the history, culture, and experience of the firm [22–24].

In summary, BDA can be understood as a strategic competence to gain analytical insights from big data, which has a specific business value and cannot be analyzed by traditional approaches such as data warehousing. The strategic competence equally results from technological and organizational capabilities.

2.2 Related Work

To identify potentially relevant related work, we conducted a literature review following the guidelines of Webster and Watson [25] and vom Brocke et al. [26]. In accordance with our research question, we examined which capabilities are seen as relevant to build a BDA competence. Furthermore, we were interested in identifying prior BDA management approaches and/or maturity models to verify the research gap. We searched the literature base for articles addressing the maturity of BDA initiatives or the required capabilities. As keywords, we used three search terms (see Table 1) to query the *AIS electronic Library* and the *EBSCO Host Business Source Complete* database. To extend the scope, we also queried *Google Scholar* using a fourth search string [27].

Table 1. Results of the literature review

Database	*Search String*			*Hits*	*Relevant*
EBSCO AIS	“big data”	AND	(“maturity model” OR “capabilities”)	16	9
	“business analytics”			17	1
	“business intelligence”			54	9
Google Scholar	allintitle: capabilities OR "maturity model" "big data"			25	7
Total				**113**	**26**

We screened the titles and abstracts of the articles to sort out irrelevant articles. Additionally, we conducted forward and backward searches based on the identified literature [25]. We then interpreted the remaining articles qualitatively, excluding non-peer-reviewed and redundant articles [25]. We further excluded articles that only use one of the keywords as either the methodological approach or as supplementary information.

In particular, we found several business intelligence maturity models (e.g. [28]–[31]) as well as reviews thereof ([32], [33]). While maturity models targeting the business intelligence or business analytics domains – as antecessors of BDA [1] – are potentially relevant for BDA, BDA differs to a great extent regarding the specific skills, competencies, and, in consequence, the required capabilities [34], [35]. Business intelligence and business analytics rather is focused on storing and analyzing structured historical data that is managed in enterprise systems or data warehouses [34], [36]. The according maturity models aim at supporting this task usually by focusing on capabilities to conduct the extraction, transformation, loading, warehousing, and historic analysis of data [34]. Moreover, some models hardly provide details, making their applicability difficult (e.g., [29], [31]), focus on technological capabilities of companies (e.g., [30]), or do not specifically look at the capabilities needed at all (e.g., [28]). Other models are too specific, focusing solely on topics such as information quality [37]. While these models might also provide guidance for a BDA scenario, BDA differs from business intelligence both in technological and organizational aspects [34]. From a technological perspective, BDA differs regarding the breadth and depth of the processed data as well as regarding the types of questions answered. In particular, BDA builds upon exploration, discovery, and prediction. The experimental nature of BDA in combination with often undefined business questions frequently results in a co-location of BDA units with business units to work closely to the analyzed products and processes [2].

While business intelligence maturity models provide a good starting point for the examination of BDA capabilities, literature only provides sporadic evidence concerning the competencies that are required in such a scenario. Debortoli et al. [34] have identified several capabilities based on an analysis of job descriptions. However, such an analysis can only provide first indications as there might be differences between the capabilities that a big data analyst ideally should bring along and those that are required in practice. An MISQ special issue discusses different analytics techniques and describes, how business intelligence and analytics (BI&A) frameworks can be used to conduct BDA and where they might have to be updated [36]. While several application domains are highlighted, specific capabilities are only mentioned as an aside, though. In particular, the article does not summarize the capabilities needed to succeed in BDA.

We were only able to identify one article focusing on this specific topic [7]. It describes challenges for governmental organizations that aim at using BDA. Proposing a preliminary set of capabilities that organizations ought to have to enhance their service through the use of big data, they address a comparable research question. However, they limit their framework to governmental organizations and only propose domain-specific capability categories. We complement this research stream by investigating, which capabilities are important for private organizations such as enterprises.

3 Research Method

To contribute to the closure of the above-mentioned research gap and to provide guidance regarding the capabilities required to perform BDA, we develop a capability model. The development of the capability model is based upon the design science paradigm, which provides guidelines for the rigorous scientific construction of novel artifacts such as constructs, models, methods, or instantiations [38], [39]. To ensure the traceability of our results, we followed a structured design science process that has been proposed to support the systematic development of capability (maturity) models [40]. This process consists of the problem definition, scoping, model development, and evaluation stage.

During the process, several requirements have to be fulfilled [40]. First, the process should be conducted iteratively. In this paper, we report on the results of the first iteration. Repetitions of the above-mentioned stages will be conducted in future iterations to improve both the structuring of the model and its level of detail. Second, the iterative nature of the process shall be used to apply multiple methodologies such as literature reviews, Delphi studies, or expert interviews [40]. As sources regarding the capabilities required for BDA are still scarce in literature, we decided to begin designing our capability model based on the results of expert interviews, which we conducted to identify relevant capabilities. Third, the developed capability model ought to be compared to existing models to ensure that it indeed provides novel results. To fulfill this requirement, we relate our capability model to existing approaches in section 5. In general, it is furthermore necessary to describe the relevance of the addressed problem to document the problem statement and justify the achieved results [40], [41]. Having documented the first two aspects in the former sections of this paper, we now focus on describing and justifying the results, i.e. the developed capability model.

The design of the capability model is informed by the Work System Theory [42]. We used the holistic enterprise perspective of the Work System Theory as conceptual basis to address all relevant facets of a company that performs BDA to deliver new products/services or to improve existing ones. This includes internal (e.g., internal business units) as well as external customers, which profit from BDA. Generally, a work system is a “view of work as occurring through a purposeful system” [42, p. 91]. It consists of nine components (see Figure 1): *customers* (people receiving products and/or services from a work system); *products* (products and/or services created by a work system); *processes* (work steps to create products and/or services); *participants* (persons doing work during the processes); *information* (either used or created); *technologies* (tools and techniques); *environment* (outside factors affecting the work system); *infrastructure* (resources used by the work system); and *strategies* (goals of the work system). We structure our capability model into different competence fields accordingly.

Moreover, we used the work system components as a structural guideline when interviewing experts about the capabilities necessary to perform BDA. We hence asked for required capabilities in each of the fields. Generally, an expert is someone with privileged knowledge about the topic of interest [43]. As BDA is not a routine task in most companies yet and we nevertheless wanted to obtain knowledge from experts (i.e., people who are intensively involved with BDA), we decided to interview members of a leading consulting firm that is specialized in supporting big data initiatives. In so doing, we gained access to specialists who had significant expertise in BDA and also had worked in different application domains, thus ensuring a broad feedback. Overall, we interviewed two managing partners (MP), three partners (P), eight managing consultants (MC), two consultants (C), and a business analyst (B). We decided to conduct semi-structured face-to-face interviews, since they are considered to be the superior data collection technique for interpretive investigations [44]. The interview guideline was structured according to the recommendations of Myers and Newman [45]. In a first part, we asked for demographic information. With respect to each work system component, we then asked for potential capabilities required to carry out BDA.

To identify relevant capabilities, we performed a cross-interview analysis [46]. First, we analyzed the gathered data using open coding techniques. Doing so allowed us to identify recurrent patterns, which we thematically grouped into segments [47]. From the segments, we could then derive the capabilities that were recurrently emphasized by the experts. Furthermore, we analyzed the segments for consistent and distinctive statements to examine if the experts' perception of the capabilities was homogeneous.

4 Capability Model

Based on the interview results, we identified 34 capabilities (CAP1-CAP34), which were mentioned as important by at least by 25% the experts.

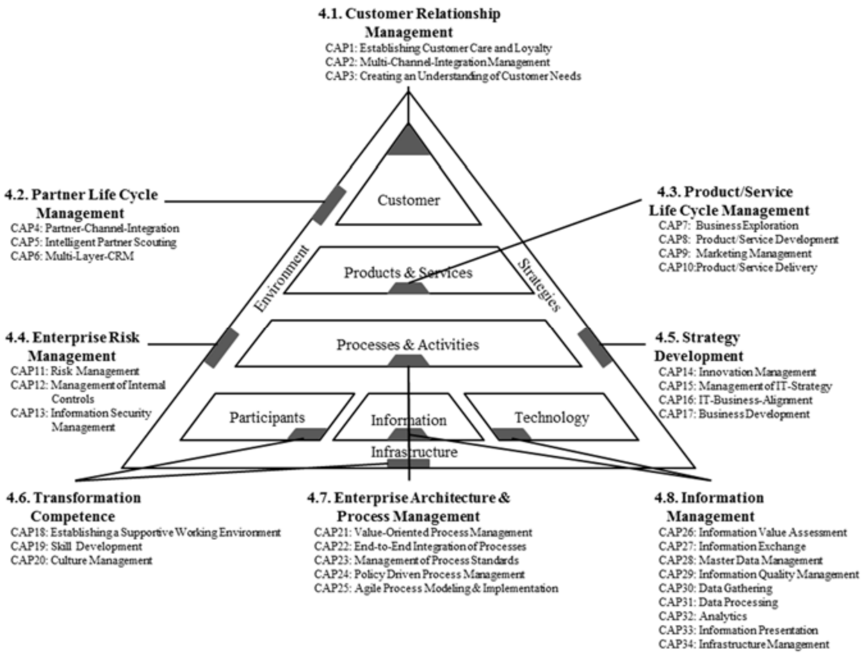


Figure 1. Big data analytics capability model according to the work system elements [42]

Based on their topic, we grouped the identified capabilities into the following eight competence fields, which each contain between two and eight capabilities: Customer Relationship Management, Partner Life Cycle Management, Product/Service Life Cycle Management, Enterprise Risk Management, Strategy Development, Transformation Competence, Enterprise Architecture and Process Management, and Information Management. We arranged the competence fields according to the elements of the Work System Theory in Figure 1. The resulting capability model provides a generic library of capabilities that can be used to assess a company's ability to successfully perform BDA.

Customer Relationship Management (CAP1-3). The capabilities in this competence field describe a company's ability to involve its customers into the value generation. The active involvement of the customers is critical to gather customer data and use it to improve or create products and services. *Establishing Customer Care and Customer Loyalty* (CAP1; mentioned by as important 6 out of 16) facilities helps involving the customer after the buying process and taking care of customer incidents, ensuring customer satisfaction and customer retention and cross-/up-selling: *"I have to make sure that my customer continuously requests my big data enabled services" (P#2). Multi-Channel-Integration Management* (CAP2; mentioned by 7 out of 16 as important) is required to connect traditional channels with new digital channels (like social media, mobile sales etc.) in order to intensify the communication with the customer and generate data: *"You have to make sure that your customers can contact you through any communication channel, which they want to use" (P#1). Creating an Understanding of Customer Needs* (CAP3; mentioned by 6 out of 16 as important) by evaluating appropriate information is a prerequisite to respond to customer requirements and to offer products/services to better suit the customers' needs: "*If I want to offer services or products, improved by big data analyses, I have to be really close to the customer needs in order to appropriately design such services and products" (P#1).*

Partner Life Cycle Management (CAP4-6). Installing BDA as a means to create or improve products/services typically requires the integration of diverse data sources including data from members of the existing value chain or new data suppliers. The capabilities contained in the competence field Partner Life Cycle Management describe a company's ability to flexibly coordinate and integrate such partners in supply chains and to create value by creating partnerships. Providing means for a systematic *Partner-Channel-Integration* (CAP4; mentioned by 6 out of 16 as important) ensures that all partners taking part in the value chain can be integrated. In particular, interfaces and communication protocols have to be established to allow the exchange of data: *"A lot of data is coming from my partners. Thus, it is very important to integrate them with the right channels" (P#1).* A company also has to have an *Intelligent Partner Scouting* (CAP5; mentioned by 4 out of 16 as important) facility to identify suitable network partners with whom to generate a competitive advantage in a specific product/service domain. In particular, this includes an examination of the credibility and reliability of the partner with respect to the dimensions time, budget, and quality to ensure that the partnership does not pose a risk: *"For topics like big data you need the right skills as well as the right partners" (MP#2).* Making use of a *Multi-Layer-CRM* (CAP6; mentioned by 4 out of 16 as important) allows sharing collected customer data with the upstream partners in the value chain, for instance by using the above-mentioned partner channels. In so doing, all members of the value chain can use the collected data of the end-customers to improve their products/services: *"I need to know the customer of my partner and to have the right communication protocols" (MC#8)*

Product/Service Life Cycle Management (CAP7-10). BDA can help to monitor the organizational environment as well as to improve products and services. To benefit from such an approach, a company has to reposition its *Product/Service Life Cycle Management* towards handling big data based products/services. Among others, this repositioning affects the *Business Exploration* (CAP7; mentioned by 4 out of 16 as important) activities of a company. On the one hand, the company needs to develop an ability to introduce new big data based products/services and the ability to withdraw them from the market if they do not add value (release often, sale early). On the other hand the company needs to be able to observe the market to discover potential new entrants: *"This is the whole topic trial and error…develop and see how the service works out and, in the next step, to improve the service" (C#1).*The repositioning also affects the *Product/Service Development* (CAP8; mentioned by 5 out of 16 as important), i.e. the ability of a company to develop big data based products/services according to the customer needs and according to the desired product/service portfolio: *"I have to consider the customers' expectations in order to deliver suitable services and products" (P #2).*By means of an adjusted *Marketing Management* (CAP9; mentioned by 4 out of 16 as important), a company has to ensure the proper marketing of big data enabled products/services. In particular, a company needs to successfully highlight the added value of such products/services, so that potential customers accept the new offerings: *"You have to ensure through specific marketing methods that a customer becomes aware of your service portfolio and says: oh, look, how awesome, they offer services based on big data analyses" (P#2).* Finally, the repositioning also affects the *Product/Service Delivery* (CAP10; mentioned by 5 out of 16 as important), i.e. the ability of a company to continuously deliver big data enabled products/services in time and with the promised quality: *"As service provider I have to be able to package and deliver continuously the insights of the big data analyses" (MC#4).*

Enterprise Risk Management (CAP11-13). The extensive usage of big data in business-critical domains implicates several potential risk factors. The category *Enterprise Risk Management* characterizes the robustness of data-driven business models and ensures an acceptable risk level through appropriate measures. A *Risk Management* (CAP11; mentioned by 5 out of 16 as important) is required to identify and prioritize data-related risks and their probability of occurrence, e.g., data theft. It includes the introduction of measures to reduce the likelihood of data risks and a monitoring system to supervise risks: *„You have to identify and prioritize both the amount of damage and the probability of occurrence of the specific risk" (MC#1).* A corporate *Management of Internal Controls* (CAP12; mentioned by 6 out of 16 as important) helps monitoring compliance guidelines through the development and maintenance of control systems (governance & compliance management system). This includes the definition of compliance guidelines with a focus on the usage of data and the culture of respecting such guidelines: *"I have to do my best in order to stay compliant and create an internal control system" (MC#1).* An *Information Security* (CAP13; mentioned by 8 out of 16 as important) concept is able to address questions of privacy and access management. It ensures the treatment and transferal of data in accordance with compliance guidelines. Access Management focuses on ensuring authentication and authorization when accessing data. Privacy Management guarantees the protection of data, specifies information ownership, and addresses digital identity issues: *"In the context of such services, information security is an essential topic because data is an asset, a production factor. If I cannot handle the data accordingly, I have a huge problem" (MP #2).*

Strategy Development (CAP14-17). BDA as veritable source of competitive advantage must also be part of a company's strategic capabilities. The capability group Strategy Development accordingly addresses the innovative usage of BDA in line with the business strategy and under the consideration of trends. This affects the *Innovation Management* (CAP14; mentioned by 4 out of 16 as important), i.e. the ability to think out of the box to develop new big data based innovations. Trends need to be detected early and monitored: *"In the context of big data I have to be able to react quickly to trends and to use them to my advantage" (MP#2).* A systematic *Management of IT-Strategy* (CAP15; mentioned by 8 out of 16 as important) ought to ensure the derivation of the IT-strategy from company's business strategy. BDA initiatives are part of the IT-strategy in order to manifest its importance and consequently its successful usage: "*In the vein of digitization this capability becomes more relevant as new business models are based on information technology" (MC#8). IT-Business-Alignment* (CAP16; mentioned by 4 out of 16 as important) is the ability to align IT-Strategy and its objectives with the objectives of the business strategy. In the context of BDA, it ensures the proper collaboration and communication between IT department, business department and, possibly, an analytics competence center in order to ensure a successful delivery of services or products: *"It has to be totally in match what IT and business is doing" (C#2). Business Development* (CAP17; mentioned by 5 out of 16 as important) includes the development and identification of new business cases, sales channels, pricing mechanisms and new business models that base on BDA analyses in order to sustain competitiveness: *"Many business models that were successful for decades are marginalized in the era of digitization" (MC#1).*

Transformation Competence (CAP18-20). BDA requires companies to transform current ways of work and their business models because of its disruptive nature. In this context, the capability group addresses a company's competence to transform and adapt to environmental dynamics. *Establishing a Supportive Working Environment* (CAP18; mentioned by 5 out of 16 as important) fosters personal responsibility, a flexible way of work and leadership supported by activity-based working structures and tools that enable employees to work efficiently and effective: *"It is all about leadership, flexible way of work, responsibility, open office and activity-based working environments and the respective supporting tools" (P#2).* For BDA approaches, it ensures a collaborative way of data sharing across departmental boundaries. *Skill Development* (CAP19; mentioned by 14 out of 16 as important) enables employees to address to new topics through trainings and workshops and, in the context of BDA, to develop the skills to handle, to structure, and to exploit heterogeneous (big) data in the business context using tools or technologies: *"Organizations nowadays do not have the people who deal with big data […] This is a big challenge to build up those capabilities […] data alone* *without the right people and the right skill sets to generate insights from this data does not add any value at the end" (C#2). Culture Management* (CAP20; mentioned by 8 out of 16 as important) encompasses introducing and maintaining a culture of open communication through appropriate methods (e.g., continuous improvement or change management). In this context, the belief that data is an essential resource needs to be part of the company's culture: *"You have to get your employees to know that data is a strategic asset on which you can capitalize. The thinking that you can actually monetize your data through services does not yet exist in companies" (C#1).*

Enterprise Architecture Management and Process Management (CAP21-26). To leverage big data to its full potential the respective technologies and tasks need to be effectively integrated into the process and IT landscape. Consequently, the capability group Enterprise Architecture Management and Process Management is about having an understanding for the interplay of applications and for the respective data streams from a process perspective as well as from a technological perspective. It compromises the following capabilities: *Value Oriented Process Management* (CAP21; mentioned by 5 out of 16 as important) is the capability to modularize the own value chain and to integrate the own value chain with the value chain of the partners. Hence, it is a prerequisite for the orchestration of end-to-end value chains across company borders and for an end-to-end data flow between partners: "*You should be able to configure your processes in a variable manner" (C#1). End-to-End Integration of Processes* (CAP22; mentioned by 6 out of 16 as important) ensures the consistency of the process architecture along the process chain over all departments and over organizational and IT boundaries and, consequently, ensures the data flow between partners: *"In the context of big data this capability is a key capability as it ensures continuous flow of data"* (MC#5). *Management of Process Standards* (CAP23; mentioned by 5 out of 16 as important) is about using and managing process standards and predefined processes. In the context of BDA, it has to guarantee the compatibility between various internal and external processes and their data exchange points: *"Against the background of collaboration it gains relevance to use process standards and to orchestrate process chains" (MC#5). Policy Driven Process Management* (CAP24; mentioned by 4 out of 16 as important) ensures that the process design and management is based on the company's legal guidelines in order to stay compliant: *"Especially in the context of big data and privacy issues, the design and the management of processes along the principles and guidelines, which ensure compliance, are a vital point" (P#1). Agile Process Modeling & Implementation* (CAP25; mentioned by 5 out of 16 as important) refers to the application of reference process frameworks, rules of modeling and compositing process chains for the design of functional processes: *"It helps to bring big data enabled products to the market more easily if you have established something like this in the company" (C#1).*

Information Management (CAP26-34). Information is the most valuable resource for BDA. Consequently, Information Management must ensure the rightful management, acquisition, processing, and distribution of information. *Information Value Assessment* (CAP26; mentioned by 9 out of 16 as important) evaluates the significance, relevance and the inherent value of (big) data: *"Data is not a by-product of value creation […] but the source of value for the company and thus for my customers" (P#2).*

To ensure the end-to-end integration of processes and the respective *Information Exchange* (CAP27; mentioned by 7 out of 16 as important) the data flow has to be established via data communication protocols and interfaces. This enables a company to perform the first task of gathering data from different data sources (e.g., partners and customers): *"It is not only about having a huge amount of data but also to have current data as this defines the value of the data" (P#1). Master Data Management* (CAP28; mentioned by 7 out of 16 as important) is about persistently saving critical company data in a structured and retrievable way: *"Is the data we have semantically correct and similar or not?" (MP#1). Information Quality Management* (CAP29; mentioned by 8 out of 16 as important) takes care of data quality and readiness (provisioning at the right time in the right quality). Furthermore, data sources and data streams are monitored and processed in order to avoid blackouts. Especially in the context of BDA the assurance of quality is a prerequisite for reasonable results as the diverse data sources and data structures change rapidly: *"If I deliver big data analytics analyses I have to make sure that my information quality is alright" (MP#1). Data Gathering* (CAP30; mentioned by 8 out of 16 as important) comprehends the identification and gathering of data and (previously unknown) information sources. Data needs to be gathered beforehand from the different sources (e.g., processes, actors, and machines) such that BDA analysis can be performed: *"I have to collect a huge data volume coming from sensors and other sources and know which sources to use" (P#2). Data Processing* (CAP31; mentioned by 8 out of 16 as important) addresses the extraction and integration of structured and unstructured, old and new, external and internal data as well as its structuring, consolidation, transformation and loading: *"If I do not have the right methods to process the data I do not even have to start at all" (C#2). Analytics* (CAP32; mentioned by 10 out of 16 as important) refers to the examination of the data in order to derive new information: *"It's about finding the value and getting information that has been unknown beforehand" (MC#4).* In this vein, new algorithms, which are capable to analyze big data, have to be developed/used that allow a company to derive insights from connected data pools. *Information Presentation* (CAP33; mentioned by 8 out of 16 as important) is essential for the digital delivery of BDA analyses. The results of those analyses needs to be visualized in a digestible manner such that the customer understands the value at a glance: *"I have to display the information I generated in a reasonable manner such that the customer can actually use the data" (MP#2). Infrastructure Management* (CAP34; mentioned by 7 out of 16 as important) ensures the provisioning and maintenance of the technological infrastructure e.g., scalable big data database systems like Hadoop: *"Which infrastructure do I need to enable big data in the first place? Do I need in-memory data bases or Hadoop clusters?" (C#2).*

5 Conclusions and Future Research

Although "big data" belongs to the most intensively discussed topics today, little research has investigated the capabilities required to effectively perform BDA. Building upon the holistic enterprise perspective of the Work System Theory and the results of 16 expert interviews, we have developed a model that provides information about several capabilities, which are potentially required to conduct BDA. In sum, we identified 34 generic capabilities that we assigned to eight capability fields: Customer Relationship Management, Partner Life Cycle Management, Product/Service Life Cycle Management, Enterprise Risk Management, Strategy Development, Transformation Competence, Enterprise Architecture and Process Management, and Information Management. The capability model gives an initial, yet unique and empirically grounded overview of the competencies that are generally required for BDA.

The results of our research have implications for academia and practice alike. From an academic perspective, the created model embodies a theory of the organizational capabilities necessary to fully leverage the potential of BDA. Although this topic appears to be of critical importance to ensure the success of big data initiatives, it hardly has been examined until now. While we have only carried out a first iteration of our design science endeavor so far, the results show that the derived capability model significantly differs from those that have been proposed for the business intelligence domain. Even though some aspects appear to be relevant in both domains, leveraging the potential of BDA appears to require additional organizational and technological capabilities. In this respect, the results of our research endeavor corroborate recent findings in literature [34]. At the same time, the developed model is both more comprehensive and more structured than other capability models that have been developed for the electronic government domain in parallel [7] or related approaches that focus on gathering concrete competencies (such as NoSQL databases, JAVA programming etc.) from job descriptions [34]. Accordingly, it advances the current body of knowledge with novel findings. For practice, the presented capability model delivers a benchmark against which companies can assess their organizational capability to leverage big data initiatives. In particular, the capability model provides the semantics for company-internal assessments of its ability to benefit from BDA. The capability model furthermore is a step towards providing an instrument, which supports the development and transformation of organizational capabilities in order to perform BDA more effectively.

Note that the relevance of the presented capabilities might vary depending on the big data scenario. This means that capabilities, which are highly important in one scenario (e.g., CAP4: Partner-Channel-Integration during supply chain improvement through BDA), might be of less importance in another scenario (e.g., in the context of a BDA service which improves customer interaction based on customer data). Moreover, some capabilities rather point to foundations for the successful usage of BDA (e.g., Infrastructure Management). In future research iterations, we will therefore also focus on separating different types of capabilities and on providing guidelines for systematically using (parts of) the model in different application scenarios. To develop a preliminary rating scheme, the analytical hierarchy process (AHP) as a proven, reproducible, and comprehensible management instrument will be used [48]. In doing so, the question, which capabilities are important in a particular usage scenario, can be solved based on pairwise comparisons of the capabilities and the input of decision makers [48]. We plan to address and illustrate the use of AHP in the context of our model in future work.

There exist several limitations in the light of which our research results have to be interpreted. Most notably, we have only conducted one iteration of the design process so far. The results might hence not be stable yet. In particular, not all capabilities are on the same level of detail and they might also not be disjunctive or exhaustive yet. Additional limitations arise because we so far have only interviewed big data consultants as experts. While this strategy originated from the impression that BDA is not a daily routine in companies yet and consultants might hence be among the persons with the most and diversified experience, our results might suffer from sample bias, because the experts worked for a single consulting company and shared certain practices. To increase the reliability of our findings, we plan to carry out additional iterations of the design process in which we will also involve experts with backgrounds from different companies. Moreover, we will apply additional research methodologies to verify our results. Finally, we have to emphasize that we have only evaluated our capability model by comparing it to related approaches so far. Although we used qualitative data for the design of the model, thus equipping it with empirical evidence, we will need to conduct empirical evaluations to further strengthen the quality of the results. To mitigate this limitation, we plan to evaluate our initial artifact, in particular its utility [49], using the four-step method of Venable et al. [50]. We intend to perform an ex-post naturalistic evaluation drawing on either a case study or a survey as evaluation method [50]. We plan to perform our evaluation at an original equipment manufacturer for the automotive industry that is on the edge of leveraging BDA to a great extent.

To fully leverage the potential of BDA, various organizational capabilities have to be developed [1], [9]. Despite the presented limitations, we hope to provide a basis to clarify and further investigate these capabilities with the presented capability model.

6 Acknowledgements

We thank Detecon International GmbH for valuable support during our first steps towards a capability model for big data analytics.

References

1. Davenport T.: Big data @ work: dispelling the myths, uncovering the opportunities. Harvard Business Review Press, Boston (2014).
2. Davenport T., Barth P., Bean R.: How ‘Big Data’ Is Different. MIT Sloan Management Review 54 (1), (2012).
3. Loebbecke C., Picot A.: Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. The Journal of Strategic Information Systems 24 (3), 149–157 (2015).
4. Agarwal R., Dhar V.: Editorial —Big Data, Data science, and Analytics. Information Systems Research 25 (3), 443–448 (2014).
5. Bhimani A.: Exploring big data’s strategic consequences. J Inf Technol 30 (1), 66–69, (2015).
6. Constantiou I. D., Kallinikos J.: New games, new rules. J Inf Technol 30 (1), 44–57 (2014).
7. Chatfield A., Reddick C., Al-Zubaidi W.: Capability Challenges in Transforming Government through Open and Big Data: Tales of Two Cities. In: Proceedings of the 36th International Conference on Information Systems, Fort Worth, USA (2015).

8. McAfee A., Brynjolfsson E., Davenport T., Patil D. J., Barton D.: Big data: The management revolution. Harvard Bus Rev 90 (10), 61–67 (2012).
9. Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C.: Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation (Accessed: 02.11.2016).
10. Mettler T., Rohner P.: Situational maturity models as instrumental artifacts for organizational design. In: Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology, Philadelphia, USA (2009).
11. Peppard J., Ward J.: Beyond strategic information systems: towards an IS capability. The Journal of Strategic Information Systems 13 (2), 167–194 (2004).
12. Tiefenbacher K., Olbrich S.: Increasing the Level of Customer Orientation-A Big Data Case Study from Insurance Industry. In: Proceedings of the 23rd European Conference on Information Systems, Münster, Germany (2015).
13. Goes P.: Editor's Comments: Big Data and IS Research. MIS Quarterly 38 (3), iii–viii (2014).
14. Yoo Y.: It is not about size. J Inf Technol 30 (1), 63–65 (2015).
15. Tan C., Sun, L., Liu K.,: Big Data Architecture for Pervasive Healthcare: A Literature Review. In: Proceedings of the 23rd European Conference on Information Systems, Münster, Germany (2015).
16. Jacobs A.: The pathologies of big data. Communications of the ACM 52 (8), 36–4.4 (2009).
17. Boyd D., Crawford K.: Critical Questions for Big Data. Information, Communication & Society 15 (5), 662–679 (2012).
18. Sharma R., Mithas S., Kankanhalli A.: Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. Eur J Inf Syst, 23 (4), 433–441 (2014).
19. Lycett M.: 'Datafication'. Eur J Inf Syst 22 (4), 381–386 (2013).
20. Duan Y., Cao G.: Understanding the Impact of Business Analytics on Innovation. In: Proceedings of the 23rd European Conference on Information Systems, Münster, Germany (2015).
21. Ulrich D., Smallwood N.: Capitalizing on capabilities. Harvard Bus Rev, 119–128 (2004).
22. Bhatt G. D., Grover V.: Types of information technology capabilities and their role in competitive advantage: An empirical study. Journal of Management Information Systems 22 (2), 253–277 (2005).
23. Teece D. J., Pisano G., Shuen A.: Dynamic capabilities and strategic management. Strategic Management Journal (18), 509–533 (1997).
24. Bharadwaj A. S., Sambamurthy V., Zmud R. W.: IT capabilities: theoretical perspectives and empirical operationalization. In: Proceedings of the 20th International Conference on Information Systems, Charlotte, USA (1999).
25. Webster J., Watson R. T.: Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly 26 (2), xiii-xxiii (2002).
26. vom Brocke J., Simons A., Niehaves B., Riemer K., Plattfaut R., Cleven A.: Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. In: Proceedings of the 17th European Conference on Information Systems, Verona, Italy (2009).
27. Rowe F.: What literature review is not. Eur J Inf Syst 23 (3), 241–255 (2014).
28. Dinter B.: The maturing of a business intelligence maturity model. In: Proceedings of the 18th Americas Conference on Information Systems, Washington, USA (2012).
29. Cosic R., Shanks G., Maynard S. B.: A business analytics capability framework. Australasian Journal of Information Systems 19, (2015).

30. Chuah M. H.: An enterprise business intelligence maturity model (EBIMM): Conceptual framework. In: Proceedings of the 5th International Conference on Digital Information Management, Thunder Bay, Canada (2010).
31. Xu P., Kim J.: Achieving Dynamic Capabilities with Business Intelligence. In: Proceedings of the 19th Pacific Asia Conference on Information Systems, Chengdu, China (2014).
32. Min-Hooi C., Kee-Luen W.: A review of business intelligence and its maturity models. African Journal of Business Management 5 (9), 3424–3428 (2011).
33. Hribar Rajterič I.: Overview of business intelligence maturity models. Management: Journal of Contemporary Management Issues 15 (1), 47–67 (2010).
34. Debortoli S., Müller O., vom Brocke J.: Comparing Business Intelligence and Big Data Skills. Bus Inf Syst Eng 6 (5), 289–300 (2014).
35. Russom P.: Big data analytics. TDWI Best Practices Report 4, 1–35 (2011).
36. Chen R., Chiang H. L., Storey V. C.: Business Intelligence and Analytics: From Big Data to Big Impact. MIS Quarterly 36 (4), 1165–1188 (2012).
37. Caballero I., Caro A., Calero C., Piattini M.: IQM3: Information Quality Management Maturity Model. Journal of Universal Computer Science 14 (22), 3658–3685 (2008).
38. Hevner A. R., March S. T., Park J., Ram S.: Design Science in Information Systems Research. MIS Quarterly 28 (1), 75–105 (2004).
39. Iivari J.: A paradigmatic analysis of information systems as a design science. Scandinavian Journal of Information Systems 19 (2), (2007).
40. Becker J., Knackstedt R., Pöppelbuß J.: Developing maturity models for IT management. Bus Inf Syst Eng 1 (3), 213–222 (2009).
41. March S. T., Smith G. F.: Design and natural science research on information technology. Decision Support Systems 15 (4), 251–266 (1995).
42. Alter S.: The work system method for understanding information systems and information systems research. Communications of the Association for Information Systems 9 (1), (2002).
43. Bogner A., Littig B., Menz W. (eds.): Interviewing experts. Palgrave Macmillan, Basingstoke, New York (2009).
44. Yin R. K.: Case study research: design and methods, SAGE Publications, Los Angeles (2014).
45. Myers M. D., Newman M.: The qualitative interview in IS research: Examining the craft. Information and Organization 17 (1), 2–26 (2007).
46. Patton M. Q.: Qualitative research & evaluation methods: integrating theory and practice. SAGE Publications, Thousand Oaks (2015).
47. Miles M. B., Huberman A. M., Saldaña J., Qualitative data analysis: a methods sourcebook. SAGE Publications, Thousand Oaks (2014).
48. Saaty T. L.: Decision making — the Analytic Hierarchy and Network Processes (AHP/ANP). J. Syst. Sci. Syst. Eng. 13 (1), 1–35 (2004).
49. Sonnenberg C., vom Brocke J.: Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research. In: Proceedings of the 7th Design Science Research in Information Systems, Las Vegas, USA (2012).
50. Venable J., Pries-Heje J., Baskerville R.: A Comprehensive Framework for Evaluation in Design Science Research. In: Proceedings of the 7th Design Science Research in Information Systems, Las Vegas, USA (2012).

The power of words: Towards a methodology for progress monitoring in design thinking projects

Mateusz Dolata[1], Falk Uebernickel[2], and Gerhard Schwabe[1]

[1] University of Zurich, Department of Informatics, Zurich, Switzerland
{dolata, schwabe}@ifi.uzh.ch
[2] University of St. Gallen, Institute of Information Management, St. Gallen, Switzerland
falk.uebernickel@unisg.ch

Abstract. The popularity of design thinking as an innovation paradigm grows continuously. More and more schools and firms implement innovation processes inspired by design thinking, but they lack easy and nonintrusive methods for monitoring the progress of teams following those processes. Consequently, interventions from coaches, teachers, or supervisors tend to rely on intuition or require intensive and intrusive examination of team dynamics. This study uses by-products from the design process and proposes automated assessment of lexical diversity as a monitoring method in process-driven design thinking projects. Thereby, it contributes to the research on the relation between text production and creativity in design projects. To the practical end, it suggests how digitalized by-products of design activities such as notes and documentation, can be leveraged to support the teams as well as coaches, teachers, and supervisors.

Keywords: design thinking, lexical diversity, documentation, creativity in teams, progress monitoring

1 Introduction

Establishing innovation culture in education and industry comes along with propagation of specific problem-solving strategies including particular tools and processes showing when and how to apply those tools. This holds also for a continuously booming innovation paradigm: design thinking (DT) – firms and schools around the world borrow from each other and propagate "successful" DT process models and implement them in their own innovation departments and classes [1–3]. They struggle, however, with an ongoing monitoring of the design teams' progress along the process [4]: Do the teams go the right path? Do they make enough improvement from one phase to another? When to intervene and when not? This paper shows the power of the words used in the documents produced by the teams: using metrics based on the vocabulary diversity bears great potential for identifying design team's progress without expensive external assessment or self-reflection episodes.

A constantly growing community of teachers, practitioners, and researchers applies DT and similar, process-dependent methods when solving wicked problems. They benefit from the agility and flexibility of the methods, as well as from the fragmentation of

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Dolata, M.; Uebernickel, F.; Schwabe, G. (2017): The power of words: Towards a methodology for progress monitoring in design thinking projects, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1156-1170

the problem and solution space enforced by the particular tools and processes. Consequently, the problems and possible solutions become more tangible, comprehensible, and user-specific. Despite large number of handbooks and guides for DT [1, 5, 6], the practitioner-oriented community still lacks deep understanding of what actually happens during a DT project. This knowledge would support the coaches and supervisors at monitoring the progress of design teams as they follow the process and at proposing meaningful interventions. As for now, they mostly rely on intuition and experience – based on the findings, we claim that their tasks could be even partially automated. Additionally, generalized set of observations regarding DT would also help the researchers with creating abstract toolset-independent models of design and creativity, which could be then transferred to novel contexts. Since the community around DT is growing continuously, we expect that the discussion on the related topics will intensify [7].

As a paradigm, DT relies on the design, production, and analysis of artifacts [1, 2]. During workshops, temporary notes capture ideas in form of easy drawings and simple descriptions for prototyping. Prototypes capture ideas in a tangible, plastic form for testing. Observations and intermediate results collected in tests are recorded as voice or text for later analysis. And, finally, results of the analysis are made persistent as text or diagrams for use in design workshops. Thanks to the recent technological progress, most of the data is shared virtually across team members, gets regularly archived in common repositories, and has textual character or format which can be automatically turned into text (e.g., voice recordings via auto-transcription or photos via character recognition). Consequently, we find ourselves in a situation where, on the one hand, organizations struggle with monitoring design teams, and on the other hand, those teams produce and file more and more potentially relevant and easy-to-process textual data. As explained below, previous research found a link between lexical diversity and the creativity of a narrative – in an educational context not related to any project group work or design documentation [8]. It also provides some evidence, that textual analysis of final documentation produced by design teams helps assessing their overall performance [9, 10]. However, as the literature leaves open whether and how those insights can be used for monitoring purposes in long-term projects, we hypothesize that *vocabulary diversity of documents produced within a DT project exhibits patterns which reflect the DT process* and, consequently, can be used to monitor the progress of design teams with regard to this process.

This study handles the above hypothesis, while analyzing a set of data produced by 15 design teams in an educational context in 6 continuous years. First, the paper discusses the process-based view on DT and ways proposed to monitor progress in similar projects; it also motivates usage of language metrics as a monitoring instrument. Second, it discusses the methodology used including the metrics used for the current analysis. Next, it provides an overview of results followed by the discussion thereof in light of what we know about DT and progress measurements. Finally, it summarizes the insights and explicates limitations of the study along with directions for future work.

2 Related Work

2.1 Design Thinking as a process

Design thinking (DT) is an innovation paradigm that gains increasing popularity and reputation. It originates from the academic world, where it was conceptualized as a teaching methodology in the area of mechanical engineering [2, 7, 11]. Later on it found its way to industry and is applied to solve practical problems in a variety of industries [5, 7, 12]. Successes of consulting firms like IDEO popularized DT as a mindset – institutions who wanted to replicate its success required, however, also a more hands-on guidance [13]. Consequently, DT was formalized as a toolset including a family of process models to be applied in previously unknown contexts [1–3, 11, 14].

Available literature proposes various formulations of the DT process. Large part of the models focus on the iteration through divergent and convergent phases when developing multiple prototypes [2, 5, 13]. More advanced models introduce blocks to differentiate between user-centered (e.g., empathize, define) and prototype-centered (e.g., ideate, prototype, test) [6]. Yet other models propose a chronological process built around phases that formulate specific requirements at the prototypes developed therein [1, 3, 11, 15]. The availability of phase-specific requirements and goals suggests that it is possible to monitor teams' progress through comparing prototypes produced in each phase with the phase goals. However, this kind of assessment requires extensive human power, especially in cases where large numbers of prototypes are produced. Also, this method is prone to subjectivity. This leads to the question whether a more objective and less extensive way of monitoring is possible.

This study focuses on the DT process model established in mechanical engineering education at Stanford [11] and adapted in a multi-national community of universities [1, 3, 16]. The macro process consists of seven milestones divided in two large phases: the divergent phase and the convergent phase. The milestones in the divergent phase include [1]: (1) *design space exploration* (DSE) – understanding the problem space and topics related to the creative challenge, (2) *critical function prototype* (CFP) – defining the principal needs of the stakeholders and addressing them through prototypes, (3) *dark horse prototype* (DHP) – reframing the challenge and questioning previously obtained insights through further "crazy" prototypes. The divergent phase closes with (4) *funky prototype* (FKY), where most promising results are compiled into complex and open prototypes. The funky milestone forms a bridge to the convergent phase consisting of the following milestones: (5) *functional prototype* (FNL) – providing a preview for the final prototype and its key aspects, (6) *X-is finished prototype* (XFP) – designing and developing a single function of the final prototype to the last extent, (7) *final prototype* (FIN) – integrating successful elements from previous milestones into single, coherent, and high-resolution prototype. Throughout the divergent phase, teams develop large numbers of low-resolution prototypes, while in the convergent phase they invest time into specifying the scope and functionality of a single prototype. Along this macro process, teams engage in multiple design iterations (micro-process) consisting of such steps as [1]: (1) need-finding and instant expertise, (2) brainstorming, (3) prototyping, (4) testing, and (5) re-defining the problem. Steps (1) and (2) have clearly a divergent character, steps (3), (4), and (5) prompt the teams to reflect and converge.

Adherence to the process is one of the essential success factors in DT projects [17]. Consequently, monitoring the progress of teams in the macro process seems to be essential for the team itself, as well as their facilitators, managers, partners, and sponsors. Even though the process provides specific prototype requirements that could be used for a kind of assessment-based progress monitoring, it would require additional human power and could lead to subjective assessment. Publications which propose particular process models leave the question open, how to monitor their implementation in an efficient and effective way [1, 5, 6].

2.2 Monitoring design process

In most general sense, monitoring involves being aware of the state of a system. With regard to DT, by monitoring we mean whether the DT-typical mindset and toolset [14] is understood and properly followed by the team. This monitoring may address interpersonal, relational, performance, or process level. The role of monitoring in DT is essential: First, it shall enable ongoing feedback regarding the facilitation success [18]. Second, it shall help with assessment of the team's performance [19]. Consequently, being aware of the team's state and its performance with regard to the process model has an essential character for a process-dependent DT implementation.

A simple measurement of creative performance may involve characterization and statistics regarding the generated ideas [20–23]. This monitoring method uses such concepts as complexity of ideas, their amount, originality, etc. While well applicable to controlled group assignments in a test setting, the adaptability of such measurements to DT seems questionable. Complex prototypes often involve several ideas, which may be new or a specification of previous ones [24]. Consequently, monitoring using statistics about ideas did not find much attention in the DT community.

Another thread of research focuses on the assessment of the social dynamics in design teams and social characteristics thereof. Two approaches dominate in this thread: psychometrics and affect-oriented measurements. Psychometric approaches rely on such tools as the Myers-Briggs Type Indicator (MBTI) [25] or similar self-reflection tools. They have been successfully employed for team composition and prediction of its outcome [26–31]. Observations of team dynamics without a self-reflection loop rely on the availability of a trained, external observer. Their potential has been confirmed

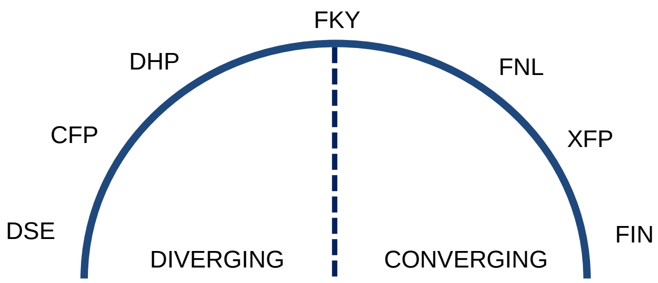


Figure 1. Phases and milestones in a DT project – chronologically from left to right [15].

with a set of studies especially from the DT community [32–35]. While team dynamics and psychometric compatibility has impact on the teams outcome, it does not help with monitoring team's progress in a process. In fact, its results may be misleading: a team that established excellent team dynamics for an early phase in the process (e.g., CFP) and does not make any progress beyond this phase, will be identified as successful by the above methods. Additionally, self-reflection and independent observation are expensive – they require dedicated time from the team or the observer. They may also be perceived as intrusive by the team. While being a well-promising idea in a laboratory, psychometric and team dynamics monitoring leaves open multiple questions on its applicability in real-world projects.

Yet another way to go involves analysis and understanding of all kinds of artifacts used or produced within creative teamwork [36–38]. This fits into the paradigm of multimodal analysis of human-generated content and actions [39]. While this approach catches the complexity of processes in artifact-oriented creative activities such as design, it has not yet turn into practical applications, thus remaining at the abstract level of frameworks and models.

To conclude, the growing popularity of DT leads to increasing attention in research. The various process models and their role for the success of DT have been extensively addressed. In parallel, several options for monitoring the team's progress with regard to the process have been proposed. However, a non-intrusive monitoring technique is still missing. Nevertheless, some preliminary studies relate language use of design teams and their overall performance [9, 10, 40, 41] – we take them as an inspiration and discuss their potential for process *monitoring*.

2.3 Language in creative processes

Language does not only describe the reality, but also "makes" it (e.g. in form of performative utterances [42]). Three primary roles of language in design can be defined: representational (explaining design process and design rationale), instrumental (communicating and recording design intent and rationale), and constructive (contributing to the design process) one [40]. Spoken language is used to describe things (ideas, concepts, information chunks) that enter the design process, while written language is used to describe things that should stay in the process or that could be useful at a later stage [10, 40]. This opens the possibility to analyze written language as a way to assess and monitor the outcome teams produce.

Language, beyond supporting and constituting design activities, is an essential element of creative processes, i.e., inventions. Vivas [43] writes that what creative persons do is "not to invent something new, but to extricate out of the subject matter at hand its own proper structure or order" – this structure is given by language in communication [8]. Several studies empirically confirm the structural nature of creativeness [44, 45]: the structure such as language drive and inform creative performance. Consequently, we argue that *language* and *creative performance* mutually influence each other – patterns discovered in one shall be reflected by specific patterns in the other one: if the creative performance is driven by the process, the language-features will be likely to reflect this process too. However, the studies that inform our argument leave it open, which characteristics of language shall be considered.

The link between creative performance and narrative performance is supported by the discourse in the area of language acquisition psychology. Narrative performance and creativity both share a number of qualities, particularly if it comes to divergent thinking [46]. Creativity, and, particularly, divergent thinking require, among others, ability to produce wide variety of ideas (flexibility) and ability to produce unusual ideas (originality). These two qualities are, also, proposed as sources for wide range of vocabulary in narrative tasks [46–48]. Lynch and Kaufman [8] identify three main factors (out of 7 main and 24 sub-factors) that have moderate factor loadings on creativity assessment: verbal diversity, readability, and consistency, whereas the first one has the highest factor loading. In other words, human judges as generally more creative perceive texts that exhibit higher diversity. Importantly, in the above studies text composition / text judgment was the main task given to the test persons. However, DT is not about writing texts, but about thriving innovative ideas. Documentation is mostly considered only a side-product of teams' activities and communication is not a goal by itself, but should support ideation and creation of the intermediate and final prototypes. Therefore, it remains open whether regularities regarding lexical diversity will hold for text others than narratives intended to be creative.

Accordingly, language of collaboration and its relation to team's creative performance has been approached within the DT community. Already in 1996, Mabogunje and Leifer [9] show that design teams in the area of mechanical engineering who introduce novel noun phrases in the mid-term and final documentation tend to produce better and more creative results than teams less innovative in their language use. Dong [10, 40] who, also, focuses on design teams in the area of mechanical engineering shows that other qualities of written language, such as coherence and presence of sentimental statements also correlate with teams' performance. However, studies leave it open how to monitor the progress of teams along a predefined process and how to use the metrics for continuous assessment rather than post-mortem analysis of projects and teams.

Overall, the role of language and evolution of language is highly relevant for creative performance and collaboration; however, it remains underspecified how language evolves over time in design teams and how it thus can be used to guide facilitation efforts. Moreover, longitudinal studies of creative performance in teams are rare, but very valuable and needed [49]. Already in 1984, McGrath [50] appealed for closer consideration and deeper research on longitudinal team collaboration. His call still remains up-to-date, in particular if it comes to teams working on creative tasks in organizational or nearly-organizational context [51–53].

3 Methodology

The context of our study is a graduate course at a Swiss University. In this class, students are taught to apply design thinking on a real-life innovation challenges, offered by industry partners. Challenges originate from different market sectors and reflect the real needs of the involved industry partners – each team works on a different challenge.

The design teams consist of graduate students with multidisciplinary backgrounds, as well as prior working experience. Each team receives coaching from DT experts. Additionally, industry partners nominate an employee as a company-internal contact partner for the team. During the course of 10 months, the teams pass the milestones described above in order to finish their projects and provide a final prototype.

The teams experience co-located and distributed collaboration. Except for the sessions with the coaches and facilitators twice a week for ca. 2 hours, the teams are free to organize their collaboration process. Sometimes, the routines get adjusted throughout the process – teams tend to intensify synchronous and co-located collaboration towards the end of the process and shortly before the particular milestones. Teams are free to choose media and technologies to support their collaboration effort. A common and non-interchangeable element for all teams forms a media wiki used as knowledge base. It was introduced into the course to enable for cross-inspiration and cross-pollution between teams within a year's class and, also, between different years' classes.

Each team regularly updates their wiki section. It includes basic information on the team, the industry partner, and the challenge. Furthermore, teams document their progress in wiki while describing (1) insights obtained through ethnographic user studies, (2) ideas implemented as prototypes along with evaluation results, and (3) experiences collected through collaboration with other teams or in research trips. They upload photos, videos, and other material into their shared space. Established in 2008, the wiki has become a large repository of design knowledge from several generations of students in this course. As of 2016 the wiki includes information for 38 challenges in German or English. For this study, we consider a subset of 15 challenges from years 2008-2015 documented in German. The overall number of documents considered equals to 630.

We decide to use the German part of the wiki because large majority of the course participants have been German or Swiss German native speakers (more than 95%). Consequently, those students who decide to run their wiki section in English, actually, choose to use a foreign language. Aware of the differences regarding lexical diversity and other related characteristics between the mother tongue and second or third language, we deliberately exclude documents written in English from this study.

To assess the lexical diversity, we employ MTLD (Measure of Textual Lexical Diversity) [54]. While, in general lexical diversity refers to ratio of total number of words to the number of different unique word stems, we decide to use the specific metric of MTLD for a particular reason – according to the empirical tests run by McCarthy and Jarvis [54] it is resistant to length-induced variance in texts of 100 to 2000 words. 98% of documents considered in this study are of this length. As most measurements for lexical diversity, MTLD takes Types-Token Ratio (TTR) into consideration, however it expresses the mean length of sequential word strings with a given TTR value (the optimal value of this factor was empirically estimated to equal to 0.7 [54]). In other words, it provides information on the average text length (measured by number of words) with a TTR above the used treshold – if the results is 100, it means that on average after each 100th word the TTS value dropped below the threshold of 0.7. To assure the measurement to be robust, the MTLD algorithm applied in this study does a forward and backward run through the text to compute the final outcome [55].

4 Results

Comparison of lexical diversity of documents from the divergent phase (milestones: DSE, CFP, DHP) and the convergent phase (FNL, XFP, FIN) shows a clear picture: while mean lexical diversity in the divergent phase equals to 242 (Standard Deviation SD = 67, Standard Error Mean SE = 3), in the convergent phase MTLD yields on average an outcome of 192 (SD = 78, SE = 10). The difference is significant and accounts for $p < 0.001$ ($t = 4.73$, $df = 73$; in a two-tailed unpaired Student's t-test without assuming equal variances – this test was chosen due to unequal samples – see below).

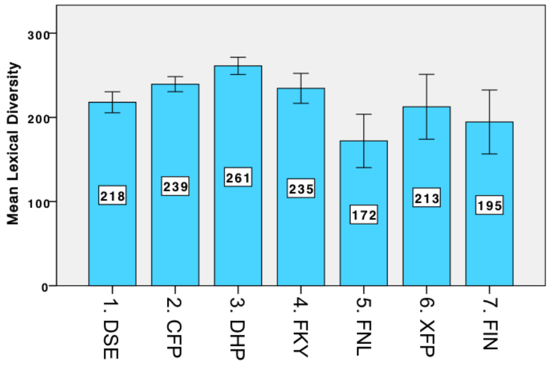


Figure 2. Mean lexical diversity (measured with MTLD) of documents for each milestone in the DT process. Data considered for 15 projects (630 documents). Error bars: 95% CI

Comparison of lexical diversity for subsequent milestones allows for analysis of process-typical tendencies. For easier comprehension, Figure 2 includes the chart representing the mean values along with error bars. In the divergent phase, we can identify a growing tendency – lexical diversity increases significantly from milestone to milestone. MTLD algorithm yields the following values: for documents in DSE it yields 218 (SD = 66, SE = 6), in CFP – 239 (SD = 63, SE = 5), and in DHP – 261 (SD = 69, SE = 5). For each step, we compute a Student's t-test and provide its results in Table 1.

FKY is the first milestone with a drop in the lexical diversity: mean lexical diversity of documents equals to 235 (SD = 81, SE = 9). However, the largest drop occurs afterwards, i.e., when the mode changes to converging. MLTD computation for the milestones in the convergent phase average to the following values: FNL – 172 (SD = 71, SE = 15), XFP – 213 (SD = 84, SE = 18), FIN – 195 (SD = 19, SE = 18). As presented in Table 1, no significant difference between those milestones could be identified – while this may be an effect of smaller number of documents considered for the statistical analysis (cf. Table 1; teams produce fewer documents as they converge and exhibit tendency to collect information in few central documents), we can definitely see that none of the phases in the convergent phase exhibit higher lexical diversity than any of the phases in divergent mode. Additionally, no clear tendency – neither growing nor lessening – can be identified for the convergent phase. We attribute the difference regarding the standard error (reflected by the length of error bars) to the number of documents teams produce in particular phases: in the divergent phase, the number of distinct documents is higher, while in the convergent phase the number of distinct documents is lower; consequently, the average figures on document lengths for each phase in each team lie further away from each other in the divergent phase.

Milestones and their short forms
Design Space Exploration DSE (111)
Critical Function Prototype CFP (193)
Dark Horse Prototype DHP (174)
Funky Prototype FKY (83)
Functional Prototype FNL (22)
X-is Finished Prototype XFP (21)
Final Prototype FIN (19)

Considered milestones	Results of a T-test regarding comparison of mean lexical diversity for the two considered milestones		
	t	df	p
DSE → CFP	2.744	220	.006 *
CFP → DHP	3.154	353	.002 *
DHP → FKY	2.579	139	.011 *
FKY → FNL	- 3.53	36	.001 *
FNL → XFP	1.691	39	.100
XFP → FIN	- 0.69	38	.491

Table 1. Results of the t-test for lexical diversity in subsequent milestones. On the left: list of all milestones with their respective short forms and number of documents considered.

To confirm the significance of the results that can be clearly seen in Figure 2 (by comparing the error bars), we decided to run a Student's t-test. Due to different number of documents in each phase in each team and because we did not want to exclude any documents from the analysis, we decided to employ a two-tailed unpaired t-test without assuming equal variances – the results for each subsequent pair of milestones can be seen in Table 1. We also run a two-tailed paired t-test where we used samples of 19 random documents from each milestone to assure that measured effects do not occur simply due to the large number of documents or difference between milestones in this regard. The results of those t-tests confirmed the significance of the results while yielding p-values $p<0.05$ in large majority of cases for milestone pairs marked with an asterisk in Table 1, i.e., those where difference regarding lexical diversity is significant.

5 Discussion

The results presented above confirm the hypothesis expressed in the introduction: lexical diversity of texts produced in a DT project can deal as a metric to monitor the progress of a team with regard to the DT process. The following reiterates this finding while pointing to the consequences it has for the monitoring of DT activities.

5.1 Particularities of the DT process

We observe, that the lexical diversity of the documents grows significantly between the milestones in the divergent phase – and reaches its peak in the dark horse phase, where teams are encouraged to leave behind previous ideas and consider totally new ideas and approaches [1, 2]. Dark horse prototypes enjoy a special status in the design thinking community – they are often presented as the source of final ideas and solutions [17]. With the above results we confirm that this milestone is, in fact, special – also with regard to the lexical diversity: it is clearly the milestone that exhibits the significantly highest lexical diversity out of all milestones. Interestingly, as soon as first consolidation efforts are made – in the funky milestone, where teams combine knowledge from previous milestones without the necessity to focus on a single coherent prototype – a slight drop in the lexical diversity occurs. We claim, that this reflects the fact, that they drop some ideas and concepts and tend to combine the most promising ones.

However, the most dramatic drop in lexical diversity occurs, where teams are pushed to converge while focusing on a single prototype built around coherent concepts. When considering the role of written language in DT process as a medium to retain things for later reference (representational and instrumental-recording role [40]), the changes in lexical diversity reflect the evolution of solution space. While in the divergent phase, the scope of potential solutions is rather large and should grow even further, in the convergent phase it is clearly limited. It is the time, where teams change their work modus from creativity to execution [1, 11] and they assess their own performance rather in terms of completing predefined tasks than in providing unexpected solutions [11]. The lack of significant differences during the convergent phase has possibly two reasons: (1) during that time, teams generate fewer documents as they concentrate on building their prototypes and do not produce new knowledge that needs saving; (2) the milestones are not as distinct as in the divergent phase – they merge into one another. Nevertheless, also in the convergent phase some variance occurs. While not being significant, the difference between FNL and XFP phases points to an interesting aspect of the DT process: the functional prototype focuses on a set of key concepts and is mostly built around a story which allows for preview of the final functionality – a successful story tends to be coherent and concise, however in XFP teams have small divergent elements, e.g., when they look for and decide between possible technical solutions to implement the single, "X" feature. Consequently, the changes of lexical diversity between particular phases reflect how goals and tasks evolve throughout a DT project.

Importantly, the projects we consider for the current study, follow a very popular DT process. Through its rigidity and the available hands-on guidance [1], it enables for a relatively easy transfer from one context to another. Consequently, it attracts attention from educational sector as well as from the industry [11]. We claim, that the presented results contribute even further to the better understanding of this particular process and its building blocks or phases. Though, it remains open, whether similar patterns can be identified in projects following another DT process model [2, 5, 6, 15].

5.2 Lexical diversity in creativity

Our findings confirm the interrelation between language production and creativity. In such complex processes as group creativity it is almost impossible to make conclusive hypothesis regarding causal relation between variety of words and divergent thinking.

Does divergent thinking cause the lexicon growth, or does lexicon growth cause divergent thinking? In accordance with previous studies [8, 42–45], we argue that language has a structural nature and, along with the predefined process, it gives a specific structure to the creative efforts of a design team [42, 43]. With this in mind, providing means to enrich the vocabulary of design teams may positively influence their creative performance in divergent phase, while reducing stimulation may positively influence team's performance in convergent phase. This opens possibilities for design of stimulation systems with specific focus on the vocabulary.

The currents study provides insights that go beyond the previous literature in twofold direction. First, it provides evidence for the link between the creative effort and textual production, even if production of the text does not constitute the central task during this effort. Consequently, it extends the findings presented by Lynch and Kaufman [8]. While our study points to the link between lexical diversity and creativity in a longitudinal project context, it leads to the question whether similar effects can be observed in local context: Do descriptions of "more creative" prototypes tend to use more diverse vocabulary? Answering this question suggest a direction for further research.

Second contribution this study makes to previous research on relation between lexicon and creativity relates more to the DT context. Mabogunje and Leifer [9] as well as Dong [10, 40] provide evidence for the relation between the overall assessment of the project (done after the project) and specific linguistics qualities of documentation (sentimentality, coherence, appearance of neologisms). We extend the list of relevant characteristics by pointing to lexical diversity as a possible and easy-to-apply metric. As opposite to previous research [9], which focused on arbitrary division of time (limited to three academic quarters) and related the results primarily with academic grades, our solutions provides higher-granularity observations on the teams' progress across the DT process. Additionally, the method proposed in here does not require extensive preprocessing [9] and can be applied to documents with a short length, which, normally, include only few noun phrases. Application of the proposed metric for performance monitoring would require further research, as discussed above.

5.3 Lexical diversity and monitoring of creative processes

Resilient methods for monitoring design processes are now necessary more than even before, as the popularity of agile design and innovation methods grows particularly fast. Using texts for assessing design teams progress provides a way to monitor DT teams without steady human control and without use of such intrusive tools as surveys and other psychometrics. In line with this, teams tend to produce and digitalize more and more data, especially side products from their design activities. This opens options going beyond the ones suggested in previous literature.

Monitoring and interventions relying on observing social dynamics of teams is extensively propagated within the DT community and is considered one of the central tasks of teams coaches [11, 17, 32–35]. Similarly, much attention is put on the controlled, conscious team composition and its effects [26–31]. Those efforts require very intensive observations, good intuition, and readiness to intervene from the coaches, thus compromising teams' autonomy and distracting them from the actual task. Even though we accept arguments for intensive coaching of innovation projects, especially with unexperienced designers, with lexical diversity we point to a source of additional, independent data useful for coaching. Since such measurement can be done independent of the team and does not require extensive human power from the coach, nor his or her intuition, it offers an objective and nonintrusive way of monitoring.

Monitoring based on statistics regarding number of singular ideas [20–23] or various qualities of prototypes [36–38] provides another way for nonintrusive process monitoring. While well applicable to particular brainstorming or prototype-design sessions, those methods are hardly applicable for DT, because prototypes and artifacts used therein are often complex and unusual. Comparing them directly with each other, even within a single project, may feel like comparing the incomparable. Under those circumstances, focusing on and monitoring side products of the design process seems reasonable. While this study provides evidence for a relevant role of vocabulary and its diversity, we claim that more metrics can be applied in similar way, just to mention argument structure in documents, cohesiveness, sentimentality, or comprehensibility.

Automatizing the method described in this study are possible directions for future design efforts to establish, e.g., a DT coach cockpit. Such a cockpit could provide on-the-go analyzes of documentation generated by the team, as well as their communication data (most teams use, e.g., WhatsApp to discuss or document ideas). The central insight from the study to be applied in such a system are the lexical diversity differences between phases. For instance, if the system detects a drop of lexical diversity in the textual material produced by a particular team between the CFP and DHP, i.e., a signal of converging tendencies in a divergent phase, the coach may intervene by providing more diverse stimulation material, introducing additional assignments, or organizing brainstorming workshops. The system may support the coach by monitoring other relevant information, such as the structure of documents or the relation between them [43, 45, 46]. Another point to observe would be the number of meetings between team members to assess whether possibilities for ideas transfer occurred at al. As for now, coaches intervene based on their subjective observation intuition, and are limited to the face-to-face meetings with the team. An independent metric, lexical diversity, makes it possible to change this punctual monitoring routine to a continuous monitoring, which in turn allows for more precise and swift interventions as well as more dedicated coaching.

6 Conclusion and Limitations

To conclude, this paper confirms the intuition inspired by earlier research in design engineering [9, 40] and in text creativity [8, 46, 47] that evolution of vocabulary in design documentation reflects the process models applied in such collaboration. Following this finding, one can speculate about the usage of documents produced by the team as a source of knowledge on team's progress. It also confirms the strong connection between the language we use and our state-of-mind. If a mutual relation between the two is assumed [43], then stimulating language will enhance creative performance, and experiencing creativity will result in observable changes in language. Consequently, both dimensions should be considered in collaborative systems designed to support creative and divergent thinking. Our findings as well as the discussion shall inspire the scientific community to reconsider the topic of creativity support, in particular, with regard to long-term innovation projects. The DT community benefits from better understanding of the process model it promotes and how it influences the communication in design teams. Finally, facilitators of design projects may use the findings to control the progress made by teams, both in organizational and educational context.

While we postulate that language can be used for monitoring and stimulation of creativity in design teams, we acknowledge the fact, that language, and particularly, lexicon is not the only important factor in this regard. We would see it as ethically doubtful, if, e.g., university grades were depending on the lexical diversity or similar measurements. In fact, linguistic performance of an individual may vary and depends on daily mood, psychophysical state, and other independent factors, such as team composition.

Additionally, this study concentrates on the differences between phases in a "normal" case, simulated by the average measurements of all teams – comparison of lexical patterns between very successful and less successful project would provide further support for monitoring. Beyond that, it would allow for performance assessment of teams and not only their adherence to the process. To conduct such study, we would, however, require extended material – dividing 15 teams in two or more groups depending on their grades results in small "n" and yields results with only limited statistical significance. We therefore call for further research in this area with a more extended data set. Also, establishing the relation between the lexical diversity and the actual creativity of the ideas described in the considered documents would provide additional evidence.

Another related question addresses the influence of mother tongue and second/third language on the linguistic performance in design process documentation. We concentrated on the German texts in this study and disclosed wiki sections written in English, i.e., in a foreign language for the authors and presumed audience. However, it would be definitely an important extension of the current study to run the same analysis on data generated by native speakers of other languages.

References

1. Uebernickel F, Brenner W, Naef T, et al (2015) Design Thinking: Das Handbuch. Frankfurter Allgemeine Buch, Frankfurt am Main, Germany.
2. Skogstad P, Leifer L (2011) A Unified Innovation Process Model for Engineering Designers and Managers. In: Meinel C, Leifer L, Plattner H (eds) Des. Think. Springer, pp 19–43
3. Uebernickel F, Brenner W (2016) Design Thinking. In: Hoffmann CP, Lennerts S, Schmitz C, et al (eds) Bus. Innov. St Galler Modell. Springer Fachmedien Wiesbaden, pp 243–265
4. Bushnell T, Steber S, Matta A, et al (2013) Using A "Dark Horse"Prototype to Manage Innovative Teams. 3rd Int. Conf. Integr. Des. Eng. Manag. Innov.
5. Brown T (2008) Design Thinking. Harv Bus Rev 86:84–92.
6. d.School (2015) Design Thinking: The d.School Bootcamp Bootleg.
7. Johansson-Sköldberg U, Woodilla J, Çetinkaya M (2013) Design Thinking: Past, Present and Possible Futures. Creat Innov Manag 22:121–146.
8. Lynch MD, Kaufman M (1974) Creativeness: Its Meaning and Measurement. J Lit Res 6:375–394.

9. Mabogunje A, Leifer LJ (1996) 210-NP: measuring the mechanical engineering design process. In: Front. Educ. Conf. 1996 FIE96 26th Annu. Conf. Proc. Of. IEEE, pp 1322–1328
10. Dong A, Hill AW, Agogino AM (2004) A Document Analysis Method for Characterizing Design Team Performance. J Mech Des 126:378.
11. Carleton T, Leifer L (2009) Stanford's ME310 Course as an Evolution of Engineering Design. Proc CIRP Des. Conf.
12. Brenner W, Witte C (2011) Business Innovation CIOs im Wettbewerb der Ideen ; mit der Methode Design Thinking. Frankfurter Allg. Buch, Frankfurt am Main
13. Brown T (2009) Change by design. HarperCollins e-books
14. Dolata M, Schwabe G (2016) Design Thinking in IS Research Projects. In: Brenner W, Uebernickel F (eds) Des. Think. Innov. Springer International Publishing, pp 67–83
15. Carleton T, Cockayne W, Tahvanainen A (2013) Playbook for strategic foresight and innovation.
16. Vetterli C, Hoffmann F, Brenner W, et al (2012) Designing innovation: Prototypes and team performance in design thinking. In: Proc. Conf. ISPIM, ISPIM, Barcelona, Spain.
17. Leifer LJ, Steinert M (2014) Dancing with Ambiguity: Causality Behavior, Design Thinking, and Triple-Loop-Learning. In: Manag. Fuzzy Front End Innov. Springer, pp 141–158
18. Mabogunje A, Leifer LJ, Levitt RE, Baudin C (1995) ME210-VDT: a managerial framework for measuring and improving design process performance. In: Front. Educ. Conf. 1995. IEEE.
19. Agogino A, Song S, Hey J (2007) Triangulation of indicators of successful student design teams. Int J Eng Educ 22:617.
20. Gero JS (1993) Modeling creativity and knowledge-based creative design. Erlbaum, Hillsdale, NJ, USA.
21. Kessel M, Kratzer J, Schultz C (2012) Psychological safety, knowledge sharing, and creative performance in healthcare teams. Creat Innov Manag 21:147–157.
22. Paulus P (2000) Groups, Teams, and Creativity: The Creative Potential of Idea-generating Groups. Appl Psychol 49:237–262. doi: 10.1111/1464-0597.00013
23. Paulus PB, Dzindolet M (2008) Social influence, creativity and innovation. Soc Influ 3:228–247.
24. Donati C, Vignoli M (2014) How tangible is your prototype? Designing the user and expert interaction. Int J Interact Des Manuf IJIDeM 1–8. doi: 10.1007/s12008-014-0232-5
25. Myers IB, McCaulley MH, Quenk NL, Hammer AL (1998) MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator. Cons. Psych. Press, Palo Alto.
26. Aronson ZH, Reilly RR, Lynn GS (2006) The impact of leader personality on new product development teamwork and performance. J Eng Technol Manag 23:221–247.
27. Barrick MR, Stewart GL, Neubert MJ, Mount MK (1998) Relating member ability and personality to work-team processes and team effectiveness. J Appl Psychol 83:377–391.
28. Bell ST (2007) Deep-level composition variables as predictors of team performance: A meta-analysis. J Appl Psychol 92:595–615.
29. 29. Kress GL, Schar M (2012) Applied Teamology: The Impact of Cognitive Style Diversity on Problem Reframing and Product Redesign Within Design Teams. In: Plattner H, Meinel C, Leifer L (eds) Des. Think. Res. Springer Berlin Heidelberg, pp 127–149
30. Kress GL, Schar M (2012) Teamology – The Art and Science of Design Team Formation. In: Plattner H, Meinel C, Leifer L (eds) Des. Think. Res. Springer, pp 189–209
31. Reilly RR, Lynn GS, Aronson ZH (2002) The role of personality in new product development team performance. J Eng Technol Manag 19:39–58.

32. Jung M, Chong J, Leifer L (2012) Group hedonic balance and pair programming performance: affective interaction dynamics as indicators of performance. In: Proc. SIGCHI Conf. Hum. Factors Comput. Syst. ACM, pp 829–838
33. Jung MF (2011) Engineering Team Performance and Emotion: Affective Interaction Dynamics as Indicators of Design Team Performance. Stanford University
34. Lande M, Sonalkar N, Jung M, et al (2012) Monitoring Design Thinking Through In-Situ Interventions. In: Plattner H, Meinel C, Leifer L (eds) Des. Think. Res. Springer, pp 211ff.
35. Sonalkar N, Jung M, Mabogunje A (2011) Emotion in Engineering Design Teams. In: Fukuda S (ed) Emot. Eng. Springer London, pp 311–326
36. Blackburn T, Swatman P (2007) Towards a framework for supporting professional teamwork: modelling human actions in small group meetings. Proc. Conf. ACIS.
37. Blackburn T, Swatman P, Vernik R (2007) Cognitive Dust: A Framework That Builds from CSCW Concepts to Provide Situated Support for Small Group Work. In: Shen W, Luo J, Lin Z, et al (eds) Comput. Support. Coop. Work Des. III. Springer, pp 1–12
38. Tan KL, Swatman P (2010) Modelling Creative Team Dynamics. In: Proc. Conf. HICSS. IEEE, pp 1–10
39. Kress G (2009) Multimodality: A Social Semiotic Approach to Contemporary Communication. Routledge, Abingdon; New York
40. Dong A (2009) The Language of Design. Springer London, London
41. Sonalkar N, Mabogunje A, Leifer L (2013) Developing a visual representation to characterize moment-to-moment concept generation in design teams. Int J Des Creat Innov 1:93–108.
42. Austin JL (1962) How to do things with words. Oxford university press
43. Vivas E (1955) Creation and Discovery Essays in Criticism and Aesthetics, Noonday Press
44. Lynch MD, Kays DJ (1967) Some Effects of Distribution of Writing Tasks and Creative Aptitude on Journalistic Performance. Journal Q 44:508–12.
45. Lynch MD, Swink E (1967) Some Effects of Priming, Incubation and Creative Aptitude on Journalism Performance. J Commun 17:372–382
46. Albert Á, Kormos J (2004) Creativity and Narrative Task Performance: An Exploratory Study. Lang Learn 54:277–310.
47. Baer J (1993) Creativity and Divergent Thinking: A Task-Specific Approach, 1 edition. Psychology Press, Hillsdale, N.J
48. Baer J (1996) The Effects of Task-Specific Divergent-Thinking Training. J Creat Behav 30:183–187.
49. Hollingshead AB, Mcgrath JE, O'Connor KM (1993) Group Task Performance and Communication Technology s. Small Group Res 24:307–333.
50. McGrath JE (1984) Groups: Interaction and performance.
51. Coughlan T, Johnson P (2009) Understanding Productive, Structural and Longitudinal Interactions in the Design of Tools for Creative Activities. In: Proc. Seventh ACM Conf. Creat. Cogn. ACM, New York, NY, USA, pp 155–164
52. Dolata M, Schwabe G (2014) Call for Action: Designing for Harmony in Creative Teams. In: Tremblay MC, VanderMeer DE, Rothenberger MA, et al (eds) Adv. Impact Des. Sci. Mov. Theory Pract., Springer, pp 273–288
53. Jehn KA, Mannix EA (2001) The Dynamic Nature of Conflict: A Longitudinal Study of Intragroup Conflict and Group Performance. Acad Manage J 44:238–251.
54. McCarthy PM, Jarvis S (2010) MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. Behav Res Methods 42:381–392.
55. Xanthos A (2014) Lingua::Diversity - measuring the diversity of text units. Avail. at CPAN

A Machine Learning Approach for Classifying Textual Data in Crowdsourcing

Marcel Rhyn[1], Ivo Blohm[1]

[1] University of St. Gallen (HSG), Institute of Information Management, St. Gallen, Switzerland
{marcel.rhyn,ivo.blohm}@unisg.ch

Abstract. Crowdsourcing represents an innovative approach that allows companies to engage a diverse network of people over the internet and use their collective creativity, expertise, or workforce for completing tasks that have previously been performed by dedicated employees or contractors. However, the process of reviewing and filtering the large amount of solutions, ideas, or feedback submitted by a crowd is a latent challenge. Identifying valuable inputs and separating them from low quality contributions that cannot be used by the companies is time-consuming and cost-intensive. In this study, we build upon the principles of text mining and machine learning to partially automatize this process. Our results show that it is possible to explain and predict the quality of crowdsourced contributions based on a set of textual features. We use these textual features to train and evaluate a classification algorithm capable of automatically filtering textual contributions in crowdsourcing.

Keywords: Crowdsourcing, Machine Learning, Text Mining, Automatization

1 Introduction

In recent years, crowdsourcing has increasingly gained attention as an innovative approach to harness the collective resources of a broad and diverse network of people over the internet. The fundamental idea of crowdsourcing is that an organization proposes the voluntary undertaking of a task to an independent group of contributors in an open call [1, 2]. It seeks to mobilize the creativity, knowledge, or distributed workforce of a large panel of people who perform value creation activities that have previously been carried out by designated agents, such as employees or third-party contractors. The approach grants scalable access to remote resources and allows tasks to be completed in a parallelized fashion regardless of time and location. In this vein, crowdsourcing has been found to greatly improve the efficiency and effectiveness of problem-solving in organizations [3, 4].

However, the potential that arises from the decentralized contributions provided by a crowd comes with a critical challenge. The quantity and complexity of information that needs to be processed and evaluated in crowdsourcing is high – especially when the contributions are submitted in a raw, textual format. In 2006, for example, more than 140’000 international participants joined the IBM Innovation Jam and submitted

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Rhyn, M.; Blohm, I. (2017): A Machine Learning Approach for Classifying Textual Data in Crowdsourcing, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1171-1185

over 46’000 ideas in a single crowdsourcing contest [5]. Similarly, the devastating earthquake in Haiti during January 2010 generated over 13’500 crowdsourced messages on online maps that were used to locate emergencies and distribute relief supplies [6]. As these contributions are submitted by a diverse network of people with different backgrounds and degrees of expertise, textual data in crowdsourcing usually entail a high amount of noise and ambiguity. Thus, the process of manually evaluating the data and filtering out low quality contributions is arduous and lengthy [2]. It generally accounts for one of the most time-consuming and cost-intensive steps in crowdsourcing [7]. For example, it took Google almost three years and 3’000 employees to condense the 150’000 proposals submitted to its *Project 10 to the 100* [2].

Text mining and machine learning algorithms represent promising solutions to cope with the vast amount of contributions in crowdsourcing [8]. They provide the means to discover patterns and extract useful information from textual data in a fast, scalable, and repeatable way [9]. In this vein, they offer the potential to automatically evaluate and filter contributions in crowdsourcing. Although multiple studies have asked for such automated approaches, research on crowdsourcing is still lacking feasible models for this task [7, 10]. Our study aims to close this gap by addressing the following research question: “What textual characteristics can be used to assess and automatically predict the quality of contributions in crowdsourcing?” To answer this question, we choose a two-pronged approach that has already been used similarly in related studies [11]. First, we apply an explanatory regression analysis to examine textual characteristics that are associated with contribution quality in crowdsourcing. Then, we use these textual characteristics for predictive modeling with machine learning algorithms. That is, we build a classifier capable of predicting the quality of the contributions based on their textual characteristics.

Hence, the contribution of our study is twofold. For researchers, we provide a set of variables and models to explain and predict contribution quality in crowdsourcing. These models and variables can be used to assess textual contributions with machine learning algorithms and, thus, contribute to a partial automatization of the evaluation process. For practitioners, we build a classifier based on the Random Forest algorithm that incorporates these variables. It is capable of automatically filtering high quality and low quality contributions submitted by a crowd and makes the process of reviewing large volumes of textual feedback more efficient.

The remainder of this paper is structured as follows. In Section 2, we discuss the characteristics of textual contributions in crowdsourcing and review existing evaluation methods for this type of data. In Section 3, we derive hypotheses regarding the relationship between textual characteristics and contribution quality in crowdsourcing. In Section 4, we describe the methodology for testing these hypotheses with a regression analysis and outline our approach for predictive modeling with machine learning algorithms. Finally, in Section 5 and 6, we analyze the results and illustrate their implications for both researchers and practitioners.

2 Related Work

2.1 Textual Data in Crowdsourcing

The fundamental principle of crowdsourcing is the use of an open call to engage a wide network of potential contributors who submit their solutions to a set of tasks broadcasted by a company [1]. In this vein, crowdsourcing facilitates the collection of information and the distribution of problem-solving to a mass of users that are coerced into productive labor [6]. On the flipside, opening up the participation to a decentralized crowd of individuals makes it more difficult to control the content and format of the data [12]. This is especially challenging for the broad range of crowdsourcing settings that are based on contributions submitted in an free text format, such as ideas on open innovation platforms [5] or user feedback in crowdsourced software testing [7]. These textual contributions represent an unstructured data format and come with several problematic characteristics regarding both their contextual and their representational quality [13]. First, there is no ground truth to contributions such as ideas, feedback, or reviews. Hence, for these types of textual contributions, it is inherently complex to assess and compare contextual characteristics such as the relevancy or the completeness of the information [14]. Members of a crowd may have different perceptions of what is relevant or interesting for such a task and will typically cover a broad range of topics in their contributions [12]. Some contributions may lack focus and specificity; others may even include contradictory or false information [2]. Second, the representation of information in textual contributions is generally of high variance and diversity [6]. Depending on their background and their degree of expertise, members of a crowd may express themselves in very distinct ways, using different expressions for similar issues or similar expressions for different issues [2]. Hence, not only is there a wide range of potential topics but also a wide range of potential descriptions for these topics. This is aggravated by the fact that textual data generated by a crowd typically entail a high amount of noise due to spelling mistakes, grammatical errors, excessive punctuation, or informal writing styles [6].

2.2 Evaluation Methods for Textual Data

Given the previously described characteristics of textual data in crowdsourcing, it is difficult to use traditional approaches to quality control [15]. For example, it is not possible to employ gold standard data as there is typically no ground truth to which the contributions can be compared. Hence, companies rely on a manual assessment of the contributions. That is, someone has to read the contributions, evaluate the quality of the content, compare it to the requirements of the task, and either accept or dismiss the input for further consideration by the company [7]. Expert panels that review and select relevant inputs represent one of the most reliable yet impractical means for this step [14]. The volume of textual data and the rate at which they are created in crowdsourcing often exceed their information processing capacities [2]. Other approaches rely on the crowd itself for the evaluation of the contributions. However, multiple studies have shown that the design of ratings scales is highly challenging and often fails to produce reliable results [16]. For example, rating scales have been found to frequently face the problems of bimodal distributions or self-selection bias [17].

In consequence, a number of studies have experimented with text mining and machine learning algorithms to support the evaluation of textual data in crowdsourcing. Walter and Back [18] use text mining algorithms to cluster ideas submitted to innovation jams in an attempt to provide decision support for expert panels reviewing the contributions. Similarly in the domain of crowdsourced software testing, existing research has used text mining approaches to automatically cluster bug reports and prioritize them for the developers [19]. In the humanitarian aid sector, Rogstadius et al. [20] and Barbier et al. [6] outline the use of text mining algorithms for clustering crowdsourced incident reports and extracting named entities (e.g., locations or family names) in order to make the coordination of appropriate responses more efficient.

Hence, existing studies have already examined how the large number of textual contributions can be clustered and organized for companies trying to analyze the multitude of diverse topics and content submitted by the crowd. We extend this body of literature by analyzing textual characteristics that can be used to explain and predict the quality of the contributions. This allows companies not only to organize the variety of contributions, but also to automatically identify relevant inputs with machine learning algorithms and filter out those that are likely not to bear any value.

3 Hypotheses Development

For developing our model, we draw upon well-established textual features discussed in related literature [21–25] to operationalize the previously described contextual and representational characteristics of crowdsourced data and examine how these features are associated with contribution quality in crowdsourcing. Contextual characteristics account for the amount and the relevancy of the information provided in textual data. Representational characteristics account for the extent to which the text is presented in a clear and intelligible manner [21].

First, the amount of information in a textual contribution has frequently been discussed as one of its most important features by related literature [22, 24, 25]. Longer contributions contain more information that could potentially be relevant for the company than shorter ones [21]. It is also easier for companies to act on feedback that is well elaborated [2], as it allows them to build a more comprehensive and coherent representation of the information in the text [14]. For example, Riedl et al. [16] note that "more accurate, understandable, and comprehensive information enables decision makers to perform better" (p. 12). On the other hand, they emphasize that contributions that are short and less elaborated tend to deliver less information that could be required for an accurate understanding of the contributions and appropriate decision making [16]. Second, related literature also emphasizes the need to consider the relevancy of the information in a contribution [21, 25]. Otterbacher [21] quantifies the extent to which a product review contains terms that are statistically important across other reviews. Similarly, Weimer and Gurevych [25] use similarity features to measure the relatedness of a post to a forum topic. For crowdsourcing in particular, relevant contributions are typically characterized as containing clear and specific information for the companies to act on [2], while vague and blurry descriptions have been found to be detrimental to contribution quality [16]. Hence, we hypothesize as follows:

Hypothesis 1. The length of a textual contribution is positively associated with the quality of the contribution.

Hypothesis 2. The specificity of the terms used in a textual contribution is positively associated with the quality of the contribution.

Besides contextual characteristics accounting for the amount and the relevancy of the information, a second layer of analysis is concerned with the representational characteristics of a contribution [21, 26]. On the one hand, representational characteristics can be used as means to measure the sophistication of a contribution [21]. For example, the readability [27] is frequently used to analyze the syntactic and semantic complexity of a text [26]. In crowdsourcing, a higher readability of a contribution should enable companies to better understand the submitted content and extract relevant cues or information more easily [14]. On the other hand, representational characteristics can be broken down to purely superficial aspects, such as the extent to which a contribution respects common writing standards or reveals irregularities [11, 25]. Poorly written contributions containing spelling errors and grammatical mistakes increase the noise and ambiguity in the data [26]. Such irregularities impose a higher cognitive load on the recipient in the company and make the contributions prone to misinterpretation [14]. Hence, they are likely to be detrimental to the interpretability or clarity of crowdsourced contributions and may render the acquisition of the embedded information more difficult for companies. Thus, we define the second set of our hypotheses as follows:

Hypothesis 3. The readability of a textual contribution is positively associated with the quality of the contribution.

Hypothesis 4. The number of spelling mistakes in a textual contribution is negatively associated with the quality of the contribution.

4 Methods and Data

In order to answer our research question, we combine two independent data sources: textual contributions from a crowdsourcing project for which we apply text mining algorithms to make them eligible for statistical analysis and an expert-based baseline measure of contribution quality. This allows detailed insights into the automated classification of the contributions with machine learning algorithms.

4.1 Data Collection

For our study, we retrieved textual data from a crowdsourcing project in the field of software testing. We conducted a crowdsourced software test in cooperation with a

German-based intermediary that ranks amongst Europe's leading platforms in this domain and manages a crowd of more than 100'000 international software testers. The test was designed as a user acceptance test for a website and has been carried out in August 2015 over the course of 5 days. It consisted of open tasks that asked the testers about their opinion on positive and negative aspects of the website as well as suggestions for further improvement. This setting was chosen for several reasons. First, user acceptance tests for websites represent one of the most frequently performed types of software tests by crowdtesting platforms, as they allow companies to gather feedback from real end users of the software [7]. Second, user acceptance tests typically lead to a large amount of textual data which are especially time-consuming to evaluate by experts or developers. Third, the feedback retrieved during user acceptance tests resemble contributions in other domains, such as ideas in innovation management or reviews in product development. This allows the results of our study to be transferred to other crowdsourcing contexts and ensures their generalizability.

We received 309 contributions in a raw textual format from 104 testers who represent the target demographic of the website and who were randomly assigned to the software test by the intermediary. On average, the contributions contained 41 words with a standard deviation of 38 words. All contributions were written in English.

4.2 Expert Evaluation of Contribution Quality

As discussed previously, there is no ground truth to contributions such as ideas, feedback, or reviews. In the absence of objective measures, it is necessary to employ an expert-based baseline measure for contribution quality [14]. Therefore, we adapted the Consensual Assessment Technique for our study [28]. We asked two software experts to manually review the feedback. Both experts are involved in the development of the website for which the user acceptance test has been conducted. Thus, they are qualified to evaluate the contributions of the crowd. They independently reviewed all test reports by using the same evaluation scheme. The evaluation scheme is based on the framework proposed by Blohm et al. [14] for crowdsourcing and includes four criteria: relevance, elaboration, feasibility, and novelty. To cover these criteria, we used questions developed by Nørgaard and Hornbæk [29] who applied them analogously for assessing usability feedback in software testing. Hence, they are suitable for our study which is concerned with similar feedback to user acceptance tests. Each criterion was rated on a 5-point Likert scale. To validate the ratings of the experts, we calculated the weighted Cohen's Kappa for each criterion [30].

Table 1. Cohen's Kappa Statistics

Relevance	**Elaboration**	**Feasibility**	**Novelty**
0.78**	0.76**	0.77**	0.73**

Note: **substantial agreement, see Landis and Koch [31]

The strength of agreement as listed in **Table 1** is substantial [31] for all criteria, indicating that we have reliable quality measures. We used the mean to aggregate the expert ratings. Since we analyze contribution quality as a multidimensional construct [14], we followed past research [32–34] and calculated a composite score for contribution quality by averaging the ratings.

4.3 Variables and Measurements

We draw upon related literature [21–25] and use the textual features derived in Section 3 as variables to explain and predict the quality of the crowdsourced contributions. We use two variables (i.e., length and specificity) to account for their contextual characteristics and two variables (i.e., readability and spelling) to account for their representational characteristics.

Length. We measure the length of a contribution by counting the total number of words per contribution.

Specificity. We measure the specificity by building the sum of all TF.IDF-indices for a contribution. The TF.IDF-index represents a term weighting scheme that accounts for the importance of a particular term in the data set based on the term frequency and the inverse document frequency [35]. Generally speaking, broad and frequently used terms by the crowd (e.g., "bad" or "design") will receive lower values than more specific terms (e.g., "unintuitive" or "navigation"). For calculating these TF.IDF-indices, we follow the commonly used bag-of-words approach with a vector space model and apply standard preprocessing steps [36]. More specifically, we tokenize the contributions by breaking them up into individual terms. We apply standard transformations to the single terms, including normalization (i.e., transforming all characters to lower-case), stop word filtering (i.e., removing terms such as articles or prepositions that bear no value for the analysis) and stemming (i.e., reducing terms to their root form to avoid duplications) with the Porter stemmer [37].

Readability. We follow Ghose and Ipeirotis [11] as well as Blohm et al. [14] and measure the readability of the text by calculating the Coleman-Liau index [27] for each contribution. This index captures the complexity of the contributions by analyzing part-of-speech tags and measuring the average length of their terms and sentences. A higher index indicates a better readability for the text.

Spelling. Finally, we measure irregularities and non-conformance to writing standards by counting the number of spelling errors per contribution. In order to ensure that the spelling errors were accurately captured, we manually reviewed all 309 contributions.

5 Models and Results

5.1 Explanatory Regression Analysis

In this section, we use regression modeling to analyze whether the textual features of the contributions are associated with their quality. The length of a contribution, the specificity of the terms, the readability of the text, and the number of spelling errors represent the independent variables. The contribution quality as rated by the experts represents the dependent variable. The results are depicted in **Table 2**.

Table 2. Regression Analysis

Coefficient	Estimate	Std. Err.	t-value	p-value	
(Intercept)	2.890	0.039	74.395	< 2.2e-16	***
Length	12.091	0.783	15.451	< 2.2e-16	***
Length (poly 2)	-4.730	0.694	-6.813	5.18e-11	***
Length (poly 3)	2.930	0.710	4.124	4.82e-05	***
Specificity	1.752	0.721	2.429	0.016	*
Readability	2.333	0.708	3.297	0.001	**
Spelling	-1.847	0.814	-2.269	0.024	*

Note: ***p < 0.001; **p < 0.01; *p < 0.05

Residual Standard Error: 0.683; R-Sq. (adj.): 0.554; F(6,302): 64.8; p-value: < 2.2e-16

It shows that the length (t = 15.451; SD = 0.783; p = < 2.2e-16) and the readability (t = 3.297; SD = 0.708; p = 0.001) of a contribution are highly significant indicators for its quality. Both features are positively correlated to the quality of the contribution. Interestingly, as indicated by the polynomials, we observe a diminishing marginal utility effect associated with number of words in a contribution, which seems conceptually reasonable. Writing 55 instead of 5 words benefits the contribution more than extending it from 150 to 200 words. Regardless of this effect, our results still support H_1 which states that the length of a textual contribution is positively associated with the quality of the contribution. The model also supports H_3 and shows that the readability of a textual contribution is positively associated with the quality of the contribution. Similarly, the specificity of the terms (t = 2.429; SD = 0.721; p = 0.016) and the number of spelling mistakes (t = -2.269; SD = .814; p = 0.024) in a contribution are significant indicators for its quality. The former is positively correlated to the quality of the contribution. The latter is negatively correlated to the quality of the contribution. These results support H_2 and H_4. The model reveals a high value for R^2 and explains the quality of the contributions significantly well. We found no evidence that potential effects between the individual contributors and their contributions affect our results. We also examined the residuals and found our model to be sound. There are no signs of heteroscedasticity nor autocorrelation. The residual show to be normally distributed. We can conclude that the proposed variables explain the quality of crowdsourced contributions at statistically significant levels. We find support for our four hypotheses and will use these findings as the foundation for predictive modeling.

5.2 Predictive Modeling

Based on the previously analyzed variables, we train and evaluate a classifier that is capable of predicting the quality of the contributions and automatically filter them. A binary classification allows for a clear selection rule [2] that decides on whether the contributions fulfill the quality requirements and are thus eligible to be forwarded to the organization for further consideration or whether they are of poor quality and should be filtered out to not induce unnecessary workload. Hence, we represent the evaluation of the contributions as a classification problem. We set the threshold for separating high quality from low quality feedback to 3.5, which is comparable to previous studies conducted for product reviews [11], and labeled the contributions. As a result, 83 contributions were classified as high-quality contributions, whereas 226 contributions were classified as low-quality contributions. This distribution is consistent with findings documented in previous studies on the quality of crowdsourced contributions [11, 2]. We tested different classification algorithms and compared the performance of Logistic Regression, Naïve Bayes, k-Nearest Neighbor, Decision Trees, and Random Forest for this study. We found the Random Forest algorithm to perform substantially better in classifying the contributions compared to the other approaches – both regarding the accuracy and the receiver operating characteristic. Our findings are consistent with comparative experiments conducted for similar classification tasks [11]. Thus, we focus on the results of the Random Forest algorithm.

The Random Forest algorithm [38] builds a large number of decision trees with different combinations of the given variables. These decision trees are internally trained and evaluated using random subsets of the same data. The Random Forest model then averages the decision trees. In this vein, it reduces the variance that comes with individual decision trees, provides information about the importance of the variables for the classification, and overcomes the risk of overfitting [39].

We use 100 decision trees for our Random Forest model and set the cutoff for the model's probability estimates at the standard value of 0.5. To build and evaluate the classifier, we followed the widely used k-fold cross-validation approach with 5 folds. That is, we randomly split our data set in a stratified manner into 5 subsets. 4 subsets are used to train the classifier with the given labels. The remaining subset does not include the quality labels and is used evaluate the performance of the classifier by comparing the labels predicted by the Random Forest algorithm to the actual labels provided by the experts. We measure the accuracy, the sensitivity, the specificity[1], and the receiver operating characteristic [40]. This procedure is repeated until each split of the data set has been used to train and evaluate the classifier.

The results of the cross-validation reveal an accuracy of 80.03% on average for our Random Forest model. Thus, by only using the four variables based on our proposed textual features, the algorithm is able to automatically predict the quality of the crowdsourced contributions and correctly classify them in over 80% of the cases. The classifier shows a very high specificity of 87.73%, indicating that it performs exceptionally well at recognizing and filtering low quality contributions. As suggested by the slightly lower sensitivity measure (60.27%), it is more difficult for the algorithm to achieve a high true positive rate. The sensitivity of the classifier can be increased by adjusting the cutoff for the probability estimates. Lowering the cutoff by 20% increases the classifier's sensitivity to 75.30%. Naturally, however, this comes at the expense of reducing its specificity to 76.56%.

---

[1] It is important to note that, in this context, specificity refers to a statistical measure that describes the true negative rate.

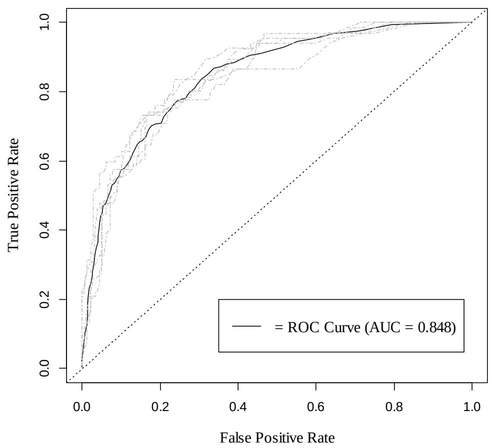


Figure 1. Receiver Operating Characteristic

The curve of the classifier’s receiver operating characteristic (ROC) is depicted in **Figure 1**. It plots the true positive rate against the false positive rate [40]. The diagonally plotted line represents the strategy of randomly guessing the quality of the contributions. A classifiers that reaches the upper triangular region of this line exploits information in the data and performs better than the random classification strategy [40]. The area under curve (AUC) is equivalent to “the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance” [40], making it also equivalent to the Wilcoxon test of ranks. Here, the AUC reveals a high value of 0.848. Hour classification algorithm performs very well.

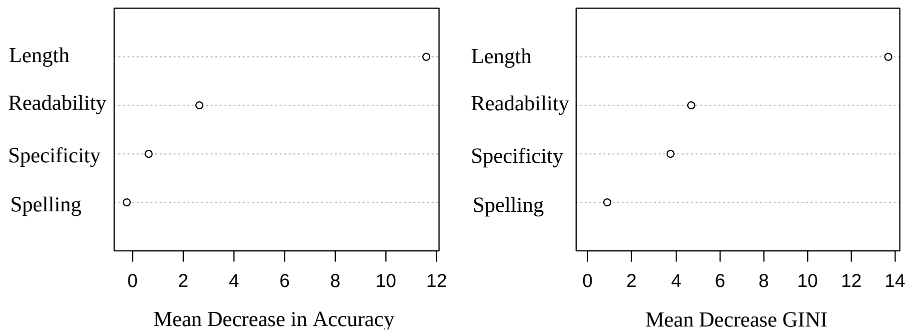


Figure 2. Variable Importance Plots

Finally, **Figure 2** displays the importance of the four proposed variables, measured by the mean decrease in accuracy and the mean decrease in node impurity (i.e., Gini index) for each variable [41]. All variables were used by the Random Forest algorithm and have predictive power. The length of the contribution is by far the most important variable for the classification. When aiming for a sparse prediction model, the variable “Spelling” may be omitted without risking much worse results.

6 Discussion and Implications

The models and results presented in the previous section yield two important findings. First, we find support for our hypotheses and show that the length of a contribution, the specificity of the terms, the readability of the text, and the number of spelling errors are all associated with contribution quality in crowdsourcing at statistically significant levels. Therefore, in a second step, we used the textual characteristics in combination with an expert-based baseline measure for contribution quality to train and evaluate an algorithm capable of predicting the contribution quality and classifying the data. Even for a small data set of 309 contributions, our Random Forest classifier achieves an accuracy of 80.03%. The algorithm has shown to perform especially well at recognizing and filtering low quality contributions. It outperforms random classification substantially and also achieves a much higher accuracy compared to a naïve classifier that would always predict the category with the majority of the ratings (i.e., 73.14%). Thus, our Random Forest algorithm proves to be very reliable. These findings have valuable implications for both researchers and practitioners alike.

6.1 Theoretical Implications

To the best of our knowledge, we are the first to show that it is possible to reliably explain and predict the quality of contributions in crowdsourcing based on textual features of the data alone. We provide empirical evidence for the relationship between both contextual and representational characteristics of contributions and their quality in crowdsourcing. This indicates that well elaborated and precise solutions, ideas, or suggestions are vital for companies trying to leverage the information submitted by a crowd. Furthermore, our results suggest that companies require textual contributions to be presented in a clear and easily interpretable manner to fully benefit from them.

Moreover, we contribute a set of models and variables to operationalize these contextual and representational characteristics. The models and variables proposed in our study have been shown to work well with algorithms capable of automatically assessing and classifying textual contributions. In this vein, we provide the foundation for partially automating the evaluation of textual data in crowdsourcing, which has frequently been requested by related literature. Kittur et al. [10] emphasized that, while "quality control is improving for tasks with a closed set of possible answers, we still have few techniques for open-ended work and highly skilled tasks" (p. 7-8). The authors specifically asked for studies to analyze potential metrics and propose feasible approaches to predict output quality. In crowdsourced software testing in particular, related work has expressed the need for efficient mechanisms to assess the quality of crowdsourced contributions and automate the evaluation of the data [7]. With our study, we close this gap and extend existing research that already uses machine learning and text mining algorithms to cluster the variety of topics covered in crowdsourcing projects [6, 18–20] by providing both the appropriate variables and models for an automated evaluation of high quality and low quality contributions in the potentially large sets of textual data. Regarding the importance of different variables, we found the length of a contribution to be the most effective indicator for explaining and predicting its quality. Both the readability of the contributions and the specificity of the terms are positively associated with the quality of the contributions at highly significant levels but reveal only moderate predictive power for classification algorithms. Interestingly, spelling errors have shown to be the least important feature for the classification and may even be omitted for sparse models. Therefore, our findings may help researchers in selecting variables for predictive modeling in crowdsourcing.

6.2 Practical Implications

Our proposed Random Forest classifier allows companies to substantially reduce the amount of information that needs to be reviewed manually. It shows that the classification algorithm is capable of automatically identifying high quality contributions in large data sets and removing those that do not fulfill the quality requirements defined by the companies or platforms. In this study, we set the threshold to only include the top 30% of the contributions. Hence, the algorithm can make the evaluation of the results submitted to crowdsourcing projects much more efficient and offers both time and cost savings. It is possible to incorporate the algorithm directly as a filter mechanism on the platforms or in tools for companies retrieving data from these platforms.

We also show that the sensitivity and specificity of the Random Forest algorithm can be adjusted to fit the preferences of practitioners. As both measures are inherently linked to each other, the decision to increase one measure will always come with the trade-off of decreasing the other. If the costs of wrongfully rejecting a high quality contribution is higher than the cost of wrongfully including a low quality contribution in the evaluation process, this is a trade-off that should potentially be considered.

Finally, our automated machine learning and text mining approach also contributes to practitioners in the domain of software testing. Related work already proposes algorithms that can be used to evaluate technical bug reports more efficiently. For example, it is possible to automatically assess the severity of the bug reports [42] and detect duplicates in the data sets [43]. As our data stem from crowdsourced software testing, we extend these findings and provide developers with an approach to facilitate the evaluation of test reports obtained in user acceptance testing, user experience testing, or usability testing. These contributions are typically submitted in a free text format and entail a high workload for the developers [7]. Our proposed classifier may help developers in evaluating these types of test reports more efficiently.

6.3 Limitations and Future Research

As with any research, our work does have its limitations. First, the manually assigned quality labels used for our data set are inherently dependent upon the rating scales and the subjective judgements of the experts. We attempted to address this issue by using scales that have been developed specifically for crowdsourced contributions as analyzed in this study [14]. Furthermore, we let two experts independently review the contributions. The Cohen’s Kappas indicate an intersubjective agreement between the experts. Second, the data set stems from a crowdsourcing project in the field of software testing. We aimed to provide as much generalizability as possible by choosing a user acceptance test setting that yields contributions similar to other crowdsourcing contexts that are based on textual data, such as ideas, feedback, or reviews.

The findings presented in this study may encourage future efforts to analyze the performance of the proposed features or models in different crowdsourcing settings and expand on our initial results. There is still great potential in making the algorithms cost-sensitive and studying the optimal trade-off between sensitivity and specificity in crowdsourcing. Furthermore, as we focused on the textual characteristics of a contribution, future work may also examine the role of non-textual characteristics and analyze features such as the experience or the expertise of the individuals who submitted the contributions. Finally, text mining and machine learning methods benefit from large data sets. Hence, we need scalable concepts for labeling crowdsourced contributions and training algorithms with more data. Addressing these issues would pave the way for leveraging the full potential of machine learning in crowdsourcing.

7 Conclusion

The process of manually reviewing and filtering large volumes of textual contributions has been a longstanding challenge in crowdsourcing. Given the unstructured format of textual data and the diversity of inputs submitted by a crowd, identifying valuable inputs and separating them from low quality contributions that cannot be used by the companies is very time-consuming and cost-intensive. In this study, we propose an approach based on the principles of text mining and machine learning to partially automatize this process. Our results indicate that it is possible to explain the quality of crowdsourced contributions purely based on textual features, such as the length of a contribution, the specificity of the words, the readability of the text, and the number of spelling errors. We use these textual features in combination with an expert-based baseline measure to train and evaluate a classification algorithm that is capable of reliably predicting the quality of the contributions and automatically filtering them for companies.

References

1. Howe, J.: The Rise of Crowdsourcing. Wired Mag. 14, 1–5 (2006).
2. Blohm, I., Leimeister, J.M., Krcmar, H.: Crowdsourcing: How to Benefit from (Too) Many Great Ideas. MIS Q. Exec. 12, 199–211 (2013).
3. Afuah, A., Tucci, C.L.: Crowdsourcing as a Solution to Distance Search. Acad. Manag. Rev. 37, 355–375 (2012).
4. Jeppesen, L.B., Lakhani, K.R.: Marginality and Problem-Solving Effectiveness in Broadcast Search. Organ. Sci. 21, 1016–1033 (2010).
5. Leimeister, J.M., Huber, M., Bretschneider, U., Krcmar, H.: Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition. J. Manag. Inf. Syst. 26, 197–224 (2009).
6. Barbier, G., Zafarani, R., Gao, H., Fung, G., Liu, H.: Maximizing Benefits from Crowdsourced Data. Comput. Math. Organ. Theory. 18, 257–279 (2012).

7. Zogaj, S., Bretschneider, U., Leimeister, J.M.: Managing Crowdsourced Software Testing: A Case Study Based Insight on the Challenges of a Crowdsourcing Intermediary. J. Bus. Econ. 84, 375–405 (2014).
8. Chen, H., Chaing, R.H.L., Storey, V.C.: Business Intelligence and Analytics: From Big Data to Big Impact. MIS Q. 36, 1165–1188 (2012).
9. Debortoli, S., Müller, O., Junglas, I.A., vom Brocke, J.: Text Mining for Information Systems Researchers: An Annotated Tutorial. Commun. AIS. 1–30 (2016).
10. Kittur, A., Nickerson, J. V., Bernstein, Michael, S., Gerber, E.M., Shaw, A., Zimmermann, J., Lease, M., Horton, J.J.: The Future of Crowd Work. In: Proceedings of the 16th ACM Conference on Computer Supported Cooperative Work, CSCW 2013. pp. 1–17. ACM, San Antonio (2013).
11. Ghose, A., Ipeirotis, P.G.: Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics. IEEE Trans. Knowl. Data Eng. 23, 1498–1512 (2011).
12. Lukyanenko, R., Parsons, J., Wiersma, Y.F.: The IQ of the Crowd: Understanding and Improving Information Quality in Structured User-Generated Content. Inf. Syst. Res. 25, 669–689 (2014).
13. Wang, R.Y., Strong, D.M.: Beyond Accuracy: What Data Quality Means to Data Consumers. J. Manag. Inf. Syst. 12, 5–34 (1996).
14. Blohm, I., Riedl, C., Füller, J., Leimeister, J.M.: Rate or Trade? Identifying Winning Ideas in Open Idea Sourcing. Inf. Syst. Res. 27, 27–48 (2016).
15. Allahbakhsh, M., Benatallah, B., Ignjatovic, A., Motahari-Nezhad, H.R., Bertino, E., Dustdar, S.: Quality Control in Crowdsourcing Systems: Issues and Directions. IEEE Internet Comput. 17, 76–81 (2013).
16. Riedl, C., Blohm, I., Leimeister, J.M., Krcmar, H.: The Effect of Rating Scales on Decision Quality and User Attitudes in Online Innovation Communities. Int. J. Electron. Commer. 17, 7–36 (2013).
17. Ghose, A., Ipeirotis, P.G., Li, B.: Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content. Mark. Sci. 31, 493–520 (2012).
18. Walter, T.P., Back, A.: A Text Mining Approach to Evaluate Submissions to Crowdsourcing Contests. In: Proceedings of the 46th Hawaii International Conference on System Sciences, HICSS. pp. 3109–3118. IEEE, Waikoloa, Hawaii (2013).
19. Feng, Y., Chen, Z., Jones, J.A., Fang, C., Xu, B.: Test Report Prioritization to Assist Crowdsourced Testing. In: Proceedings of the 10th Joint Meeting on Foundations of Software Engineering, ESEC/FSE 2015. pp. 225–236. ACM, Lombardy (2015).
20. Rogstadius, J., Vukovic, M., Teixeira, C.A., Kostakos, V., Karapanos, E., Laredo, J.A.: CrisisTracker: Crowdsourced Social Media Curation for Disaster Awareness. IBM J. Res. Dev. 57, 1–13 (2013).
21. Otterbacher, J.: “Helpfulness” in Online Communities: A Measure of Message Quality. In: Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI ’09). pp. 955–964. ACM, Boston (2009).
22. Jeon, J., Croft, W.B., Lee, J.H., Park, S.: A Framework to Predict the Quality of Answers with Non-Textual Features. In: Proceedings of the 29th Annual International ACM SIGIR Conference. pp. 228–235. ACM, Seattle, Washington (2006).
23. Liu, J., Cao, Y., Lin, C.-Y., Huang, Y., Zhou, M.: Low-Quality Product Review Detection in Opinion Summarization. In: Proceedings of the 2007 Join Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. pp. 334–342. , Prague (2007).

24. Kim, S.-M., Pantel, P., Chklovski, T., Pennacchiotti, M.: Automatically Assessing Review Helpfulness. In: Proceedings of the 2006 Conference of Empirical Methods in Natural Language Processing (EMNLP 2006). pp. 423–430. , Sydney (2006).
25. Weimer, M., Gurevych, I.: Predicting the Perceived Quality of Web Forum Posts. In: Proceedings of the 2007 Conference on Recent Advances in Natural Language Processing (RANLP). pp. 643–648. , Borovets, Bulgaria (2007).
26. Agichtein, E., Castillo, C., Donato, D., Gionis, A., Mishne, G.: Finding High-Quality Content in Social Media. In: Proceedings of the 2008 International Conference on Web Search and Data Mining. pp. 183–193. ACM, Palo Alto (2008).
27. Coleman, M., Liau, T.L.: A Computer Readability Formula Designed for Machine Scoring. J. Appl. Psychol. 60, 283–284 (1975).
28. Amabile, T.M.: Social Psychology of Creativity: A Consensual Assessment Technique. J. Pers. Soc. Psychol. 43, 997–1013 (1982).
29. Nørgaard, M., Hornbæk, K.: Exploring the Value of Usability Feedback Formats. Int. J. Hum. Comput. Interact. 25, 49–74 (2009).
30. Cohen, J.: Weighted Kappa: Nominal Scale Agreement with Provision for Scaled Disagreement or Partial Credit. Psychol. Bull. 70, 213–220 (1968).
31. Landis, J.R., Koch, G.G.: The Measurement of Observer Agreement for Categorical Data. Biometrics. 33, 159–174 (1977).
32. Barki, H., Pinsonneault, a.: Small Group Brainstorming and Idea Quality: Is Electronic Brainstorming the Most Effective Approach? Small Gr. Res. 32, 158–205 (2001).
33. Blohm, I., Bretschneider, U., Leimeister, J.M., Krcmar, H.: Does Collaboration Among Participants Lead to Better Ideas in IT-based Idea Competitions? An Empirical Investigation. In: Proceedings of the 43rd Annual Hawaii International Conference on System Sciences, HICSS 2010. pp. 1–10. IEEE, Honolulu (2010).
34. Gallupe, R.B., Dennis, A.R., Cooper, W.H., Valacich, J.S., Lana, M., Nunamaker, J.F.: Electronic Brainstorming and Group Size. Acad. Manag. J. 35, 350–369 (1992).
35. Hotho, A., Nürnberger, A., Paaß, G.: A Brief Survey of Text Mining. LDV Forum - Gld. J. Comput. Linguist. Lang. Technol. 20, 19–62 (2005).
36. Feldman, R., Sanger, J.: The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data. Cambridge University Press, Cambridge (2007).
37. Porter, M.F.: An Algorithm for Suffix Stripping. Program. 14, 130–137 (1980).
38. Breiman, L.: Random Forests. Mach. Learn. 45, 5–32 (2001).
39. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York (2009).
40. Fawcett, T.: An Introduction to ROC Analysis. Pattern Recognit. Lett. 27, 861–874 (2006).
41. Liaw, A.: Package "randomForest," https://cran.r-project.org/web/packages/randomForest/randomForest.pdf.
42. Lamkanfi, A., Demeyer, S., Soetens, Q.D., Verdonckz, T.: Comparing Mining Algorithms for Predicting the Severity of a Reported Bug. In: 15th European Conference on Software Maintenance and Reengineering, CSMR 2011. pp. 249–258. IEEE, Oldenburg (2011).
43. Runeson, P., Alexandersson, M., Nyholm, O.: Detection of Duplicate Defect Reports Using Natural Language Processing. In: Proceedings of the 29th International Conference on Software Engineering, ICSE 2007. pp. 499–508. IEEE, Minneapolis (2007).

Erhebung und Diagnostik von Markenassoziationsnetzwerken auf Grundlage nutzergenerierter Inhalte

Marc Egger[1], Gloria Volkmann[1], und Detlef Schoder[1]

[1] Universität zu Köln, Seminar für Wirtschaftsinformatik und Informationsmanagement, Köln, Deutschland
{egger,volkmann,schoder}@wim.uni-koeln.de

Abstract. Bei Konsumentscheidungen greifen Verbraucher auf ihr vorhandenes Wissen über Marken zurück, welches sie im Gedächtnis in Form von Markenassoziationsnetzwerken organisieren und deren Kenntnis es erlaubt Verbraucherentscheidungen besser zu verstehen. Traditionell werden Markenassoziationsnetzwerke mittels reaktiver, befragungsorientierter Instrumente direkt von Verbrauchern erhoben. Gleichzeitig hinterlassen unzählige Verbraucher online in nutzergenerierten Inhalten (UGC) ihre verbal explizierten Wahrnehmungen freiwillig und unangeleitet und erzeugen damit eine neue Datengrundlage zur nichtreaktiven, passiven Erhebung von Markenassoziationsnetzwerken. Traditionelle inhaltanalytische Verfahren, wie die qualitative Inhaltsanalyse, ermöglichen aufgrund der Quantität, Heterogenität und Erstellungsgeschwindigkeit von UGC entweder nur Stichprobenbetrachtungen, oder sind mit prohibitiv hohen Aufwänden verbunden. Somit erfordert die nichtreaktive Erhebung von Markenassoziationsnetzwerken aus UGC neue, computergestützte Verfahren, zu denen der vorliegende Artikel einen methodischen Beitrag leistet. Auf Grundlage von sprachlichen Mustern traditionell befragungsorientiert erhobener Markenassoziationen wird ein Text Mining Verfahren zur Extraktion von Markenassoziationsnetzwerken aus UGC vorgeschlagen, anhand einer Vergleichsstudie diskutiert und der Einsatz für die Markenführung demonstriert.

Keywords: Association Networks, Brand Diagnostics, Text Mining, Network Analysis

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Egger, M.; Volkmann, G.; Schoder, D. (2017): Erhebung und Diagnostik von Markenassoziationsnetzwerken auf Grundlage nutzergenerierter Inhalte, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1186-1200

1 Einleitung

In Zeiten der Informationsüberlastung und sich zunehmend angleichenden Produkten und Services können Verbraucher diese oft nur schwer unterscheiden [1, 2]. Bei Unsicherheit, sowie in Konsumentscheidungssituationen greifen sie daher auf ihr bereits vorhandenes Wissen über Marken als Entscheidungshilfe zurück [1, 3-5]. Dieser als Markenwissen bezeichnete Beeinflussungsfaktor wirkt auf den kundenorientierten Markenwert [3, 6, 7], wenn Verbraucher mehr oder weniger positiv auf die Elemente des Marketing-Mixes eines markierten- gegenüber eines unmarkierten Produktes oder Services reagieren [1]. Der kundenorientierte Markenwert liefert somit diagnostische Informationen für das Entstehen des monetär messbaren, finanzorientierten Markenwertes, welcher die vom Verbraucher akzeptierte Wertdifferenz zwischen einem unmarkierten und einem markierten aber gleichwertigen Produkt oder Service bezeichnet. Um sich Wettbewerbsvorteile zu verschaffen und den finanzorientierten Markenwert zu erhöhen, müssen Unternehmen daher Strategien zur Steuerung des kundenorientierten Markenwerts entwickeln [1, 2]. Die Ableitung von Strategien setzt allerdings zunächst eine entsprechende Konzeptualisierung sowie die Operationalisierung entsprechender Messgrößen voraus. Gemäß der Konzeptualisierung nach Keller [1] wird der kundenorientierte Markenwert durch das Markenwissen im Gedächtnis des Verbrauchers bestimmt. Dieses setzt sich aus der Markenbekanntheit (Brand Awareness) und dem Markenimage (Brand Image) zusammen [1]. Während die Markenbekanntheit die (Wieder-)erkennungs- und Erinnerungsleistung an die Marke beeinflusst, manifestiert sich das Markenimage in Netzwerken von Markenassoziationen im Gedächtnis der Verbraucher [1]. Diese netzwerkorientiere Sichtweise des Markenimages anhand Markenassoziationsnetzwerken gründet auf der Theorie des Human Associative Memory Models [8] und der Spreading Activation Theorie [4] aus der Kognitionspsychologie. Demnach wird das semantische Gedächtnis des Menschen als Netzwerk aus informationstragenden Knoten (Assoziationen) sowie Kanten unterschiedlicher Stärke verstanden und Aktivierungen einzelner Knoten pflanzen sich entlang der Kanten fort, um weitere Knoten zu aktivieren. Markenassoziationen repräsentieren solche informationstragenden Knoten und entstehen aufgrund direkter (z.B. Verwendung) wie auch indirekter Erfahrungen (z.B. Word-of-Mouth (WoM) der Konsumenten mit dem Leitungsangebot eines Unternehmens [1, 2, 9]. Ein netzwerkorientierter Ansatz Markenassoziationen zu repräsentieren, soll letztlich dazu beitragen, ein besseres Verständnis davon zu erhalten, wie Verbraucher die Marke- sowie das sich aus Menschen, Orten und Nutzungsszenarien zusammensetzende Umfeld der Marke wahrnehmen [10]. Dies schafft eine Grundlage, Präferenzen und Auswahlentscheidungen, und damit die Determinanten des kundenorientierten Markenwertes zu verstehen [1, 10, 11]. Henderson et al. [10, 11] zeigen diese diagnostische Stärke, indem sie aus der Struktur von Markenassoziationsnetzwerken Markenmechanismen wie Markengleichheit, -Verwechselbarkeit oder Verwässerungseffekten erklären können. Keller [1] argumentiert zudem, dass kundenorientierter Markenwert entsteht, wenn Verbraucher starke, vorteilhafte, sowie einzigartige Markenassoziationen besitzen. Letztlich setzen strategische Maßnahmenplanungen zur Schaffung von kundenorientiertem Markenwert jedoch zunächst die Erhebung von Markenassoziationsnetzwerken als diagnostische Kenngröße voraus. Klassischerweise erfolgt die Erhebung von Markenassoziationen zur Markenimagemessung anhand befragungsorientierter Erhebungsinstrumente zu denen Farsky [12] einen umfassenden Überblick gewährt. Den Erhebungsmethoden ist gemein, dass sie auf reaktiven Befragungen basieren. Aufgrund von Aufwand und Kosten sind Unternehmen jedoch zunehmend an neuen kosteneffizienten Methoden interessiert [13]. Gleichzeitig nutzen heutzutage unzählige Verbraucher das Internet und Social Media Kanäle um freiwillig und unangeleitet, öffentlich zugänglich ihre Einstellungen, Gefühle und Meinungen zu verschiedensten Themen zu verbreiten [14-16]. Diese öffentlichen und häufig in Textform vorliegenden nutzergenerierten Inhalte (User-generated Content : UGC) bieten als qualitative Datengrundlage das Potenzial, Explizierungen von individuellem Markenwissen zu enthalten was eine neue Datenquelle zur nichtreaktiven, passiven Erhebung von Markenassoziationsnetzwerken eröffnen würde. UGC stellt damit eine potentiell kostengünstige und zudem jederzeit verfügbare Alternative zu mittels traditioneller Verfahren erhobener Stichproben dar [13]. Da Markenassoziationen in textbasiertem UGC allerdings durch Worte und Phrasen ausgedrückt werden, müssen diese zunächst identifiziert und extrahiert werden. Als traditionelle Methode kann hier die qualitative Inhaltsanalyse in Betracht gezogen werden, bei welcher ein Analyst Markenassoziationen und Relationen zwischen diesen Assoziationen manuell identifiziert und codiert. Die Menge und Erstellungsfrequenz, sowie das dezentrale Vorkommen von UGC führt jedoch dazu, dass entweder nur eine Stichprobe des insgesamt vorhandenen UGC betrachtet werden kann, oder dass große (manuelle) Aufwände notwendig werden, die zu prohibitiv hohen Kosten der Erhebung führen können. Dies führt zur Notwendigkeit neuer computergestützter Verfahren um den öffentlich vorliegenden UGC zur Ableitung von Markenassoziationsnetzwerken zugänglich zu machen. Der vorliegende Artikel widmet sich dieser Forschungslücke und schlägt ein Verfahren vor, um computergestützt Markenassoziationsnetzwerke aus UGC zu extrahieren. In Abschnitt 2 wird hierzu zunächst ein Überblick über verwandte Arbeiten gewährt und von der vorliegenden Arbeit abgegrenzt. Darauf aufbauend schlägt Abschnitt 3 dann ein neues, automatisiertes Textanalyseverfahren zur Extraktion von Markenassoziationsnetzwerken aus UGC vor. Dieses Verfahren wird dabei induktiv anhand von Sprachmustern aus traditionell erhobenen Markenassoziationen entwickelt. Abschnitt 4 wendet die vorgeschlagene Methodik auf das Beispiel der Marke Google im Kontext des autonomen Fahrens an, und diskutiert den Einsatz der vorgeschlagenen Methodik als diagnostisches Tool für Markenführungsmaßnahmen. Abschnitt 5 schließt mit der Diskussion der Limitierungen und gibt einen Ausblick auf künftige Forschung.

2 Verwandte Arbeiten

Die Extraktion von marketingrelevanten Informationen aus UGC hat in den vergangenen Jahren an Bedeutung gewonnen. Der vorliegende Artikel der eine Methode zur Extraktion von Markenassoziationen aus UGC vorschlägt, grenzt sich wie folgt von anderen Beiträgen ab. Zum einen finden sich Beiträge bei denen Markenassoziationsnetzwerke entweder nicht aus UGC, sondern traditionell über

Befragung erhoben [17-19] werden. In anderen Beiträgen stellt UGC zwar den Untersuchungsgegenstand dar, jedoch werden anstatt Markenassoziationsnetzwerken beispielsweise Marktstrukturinformationen [16] in Netzwerken abgebildet. Weiterhin wurden zu Markenassoziationen ähnliche Konstrukte mittels Text Mining aus UGC extrahiert, jedoch keine Netzwerkstrukturen erhoben, sondern beispielsweise Ranglisten aus Markenassoziationen [20], Produkteigenschaften [21] oder allgemeinen Features [22]. Eine erste Arbeit welche die Extraktion von Markenassoziationsnetzwerken aus UGC und den Nutzen für die Markenführung zeigt [23], ist die Arbeit von Gensler et al. [23] . Im Gegensatz zum hier vorliegenden Beitrag fokussieren Gensler et al. [23] dabei jedoch auf eine spezifische Produktbewertungsplattform als eine bestimmte Ausprägung von UGC und nutzten hier besondere Eigenschaften der dort vorliegenden Produktbewertungen um Markenassoziationen zu extrahieren anstatt den Textinhalt [23]. Dies erlaubt es zwar, das diagnostische Potenzial der Erhebung von Markenassoziationsnetzwerken aus Produktbewertungen zu zeigen, wird jedoch dadurch eingeschränkt, dass die vorgestellte Methodik nicht ohne weiteres auf andere UGC Quellen erweitert werden kann. Damit kann die vorgestellte Methodik nur eine spezifische Stichprobe des verfügbaren UGC auswerten. Der vorliegende Artikel ordnet sich in die bestehende Forschung dahingehend ein, den Fokus von Gensler et al. [23] unter Anwendung zu [22] oder [21] ähnlicher Textanalyseverfahren zu erweitern und damit erstmalig einen Vorschlag zur ganzheitlichen Extraktion von Markenassoziationsnetzwerken aus UGC zu unterbreiten.

3 Vorgeschlagenes Verfahren zur Extraktion von Markenassoziationsnetzwerken aus UGC

Im Folgenden wird das Verfahren zur Erhebung und Konstruktion von Markenassoziationsnetzwerken aus UGC vorgestellt und anhand des Beispiels der Marke Google im Kontext des autonomen Fahrens illustriert. Das vorgeschlagene Verfahren folgt einem 8-stufigen Erhebungsprozess, der sich an das Vorgehen von Gensler et al. [23] anlehnt. Dies sind (1) Festlegung des Untersuchungsobjektes, (2) Datensammlung, (3) Datenbereinigung, (4) Extraktion von Markenassoziationen, (5) Sentimentanalyse, (6) Bestimmung der Assoziationsstärke, (7) Bestimmung der Verbindungen zwischen Assoziationen, sowie (8) die Konstruktion des Markenassoziationsnetzwerks.

3.1 Festlegung des Untersuchungsobjektes

Im ersten Schritt ist zunächst die Entscheidung für ein Untersuchungsobjekt zu treffen. Hierbei kann es sich um eine Marke, Dachmarke, Produktmarke oder ein Themenumfeld handeln. Im Rahmen dieses Beitrags wird als illustrierende Studie, das Untersuchungsobjekt als die Marke „Google“ innerhalb des Themenfeldes „autonomes fahren“ definiert. Beim autonomen Fahren handelt es sich um eine Innovation im Automobilbereich, bei dem sich Fahrzeuge in naher Zukunft selbstständig fortbewegen ohne das Eingreifen des Fahrers. Im Rahmen dieser Technologie haben

Fahrzeughersteller wie Audi, BMW und Mercedes oder Tesla, jedoch auch vermeintlich branchenfremde Unternehmen wie Google oder Apple Milliarden von Dollar in die Entwicklung investiert. Schlussendlich ist jedoch die Akzeptanz von Verbrauchern maßgeblich für den Erfolg der Markteinführung und Verbraucher treffen ihre Entscheidung auf Grundlage ihres vorhandenen Wissens in Form von Assoziationsnetzwerken. Daher ist die Kenntnis der Assoziationsnetzwerke, insbesondere im Umfeld des autonomen Fahrens, ein wichtiger Schritt für Marken potenzielle Wirkmechanismen im Hinblick auf ihren Markteintritt zu verstehen oder um ihre Marken so zu positionieren dass ein erfolgreicher Markteintritt möglich wird.

3.2 Datensammlung

Innerhalb der Datensammlung müssen im Internet solche Texte gesammelt werden, welche von Verbrauchern veröffentlicht wurden und sich thematisch mit dem Untersuchungsobjekt befassen. Im Rahmen des hier beschriebenen Verfahrens wird hierzu die Nutzung von allgemeinen Web-Suchmaschinen wie Google.com oder Bing.com vorgeschlagen. Hierzu muss eine entsprechende boolesche Suchabfrage formuliert und an die Web-Suchmaschine geschickt werden. Die von dieser zurückgelieferten Verweise auf Webseiten (uniform ressource locator – URL), sowie die Webseiten selbst werden mittels eines sog. Web-Crawlers abgerufen und gespeichert. Im Rahmen der illustrierenden Studie wurde für diesen Prozess ein Web-Crawler entwickelt, welcher die boolesche Suchabfrage aus Abbildung 1 an die Google Suchmaschine übermittelt, und auf Grundlage der zurückgelieferten URLs die korrespondierenden HTML Webseiteninhalte in einer zentralen Datenbank abspeichert. Insgesamt wurden hierbei 1.1 Millionen englischsprachige HTML Webseiteninhalte von unterschiedlichen URLs gesammelt, welche sich über einen Zeitraum von zehn Jahren zwischen April 2005 sowie April 2015 erstrecken.

Google AND (“autonomous driving” OR “self driving car” OR “self driving cars” OR
“driverless car” OR “driverless cars” OR “autonomous vehicle” OR “automated driving”
OR “piloted driving” OR “driverless car”)

Abbildung 1: Search Query

3.3 Datenbereinigung

Um möglichst relevante UGC Textbeiträge für die Extraktion von Markenassoziationen zu verwenden, wird ein dreistufiger Bereinigungsprozess vorgeschlagen. In der ersten Stufe müssen zunächst Textbestandteile aus den jeweils gesammelten, einzelnen Webseiten erkannt und isoliert werden, da diese üblicherweise nicht nur aus den eigentlichen Textbeiträgen, sondern auch aus anderen Elementen wie z.B. Navigations-, Werbungs-, Kopf- und Fußbereichen bestehen. Diese Elemente werden zusammenfassend als Boilerplate bezeichnet und können möglicherweise Teile der booleschen Suchanfrage beinhalten, obwohl sie inhaltlich keine Informationen tragen (z.B. „autonome Autos“ als Menüpunkt in Navigationselementen). Jede einzelne Webseite kann dabei mehrere irrelevante Boilerplate-Elemente wie auch mehrere relevante Inhaltselemente beinhalten. Zur Erkennung und Entfernung von Boilerplate-Elementen wird der Ansatz von Kohlschütter [24] empfohlen und verwendet. Dieser basiert auf generischen Eigenschaften von Web-Dokumenten (z.B. link-zu-text Verhältnis) und kann daher als robust gegenüber heterogenen Inputdokumenten angesehen werden, was eine Charakteristik der hier beschriebenen Datensammlung ist. Die zweite Bereinigungsstufe dient der Erkennung und Bereinigung von Duplikaten, welche sonst möglicherweise zu Verzerrungen aufgrund von Mehrfachzählungen - beispielsweise bei der Bestimmung von Markenassoziations- und Vernetzungsstärken - führen können. Duplikate treten beispielweise dann auf, wenn der selbe Inhalt unter unterschiedlichen URLs abrufbar ist (z.B. Druckansichten oder Produktbeschreibungen). Absolute Duplikate können hierbei einfach mittels zeichenweiser Vergleiche erkannt und bereinigt werden. Im letzten Schritt der Datenbereinigung werden nun nicht nutzergenerierte Inhalte (non-UGC) von der Datengrundlage entfernt, da die im Ziel zu erhebenden Markenassoziationsnetzwerke die Verbraucherperspektive abbilden sollen und somit ausschließlich von Verbrauchern erzeugte Inhalte (UGC) von Interesse sind. UGC findet sich im Internet verstreut auf sozialen Netzwerken, Foren, Blogs, Produktbewertungsplattformen, E-Commerce Webseiten und letztlich potenziell auf jeder Webseite, die eine Kommentarfunktionalität anbietet. UGC liegt also dezentral vor, und vermischt sich auch mit non-UGC wie z.B. redaktionelle Artikel, Produktbeschreibungen oder Werbung) [25]. Somit ist es zweckmäßig in den gesammelten Dokumenten, non-UGC Dokumente zu bereinigen. Egger et al. [25] haben gezeigt, das textbasierter UGC inhärente Eigenschaften aufweist, die es erlauben ein binäres Textklassifikationsproblem zu formulieren und mittels Maschinenlernverfahren zu adressieren. Im Rahmen der illustrierenden Studie wird dem Vorgehen von Egger et al. [25] gefolgt und manuell eine Stichprobe von 422 Dokumenten als UGC (214) und non-UGC (208) annotiert um ein sog. Gold Set zu erzeugen. Dieses Gold Set stellt die Grundlage an Dokumenten, welche für das Anlernen eines Klassifizierers verwendet werden und beinhaltet für jedes dort enthaltene Dokument die korrekte Klassifizierung in UGC oder non-UGC. Auf Grundlage dieses Gold Sets wird dann ein auf Support-Vector-Machines (SVM) [26] basierender Klassifizierer unter Verwendung von 10-fold Cross-Validation angelernt. Die 10-fold Cross Validation ist dabei eine Vorgehensweise den Klassifizierer anzulernen und zu validieren. Hierbei wird das Gold-Set in 10 gleichgroße Segmente unterteilt. Von diesen Segmenten werden in 10 Iterationen, jeweils 9 Segmente zum Anlernen des Klassifizierers und eines zur Validierung des Klassifizierers verwendet. Die Validierung erfolgt anhand der Metriken Precision, Recall und Accurracy (Abbildung 2). „true positive“ entspricht hierbei einer gemäß des Gold Sets korrekten Klassifizierung als UGC, „true negative“ einer korrekten Klassifizierung als non-UGC, „false positive“ einer Fehlklassifizierung als UGC und „false negative“ einer Fehlklassifizierung als non-UGC.

$$precision = \frac{true\ positive}{true\ positive + false\ positive}$$

$$recall = \frac{true\ positive}{true\ positive + false\ negative}$$

$$accurracy = \frac{true\ positive + true\ negative}{true\ positive + true\ negative + false\ positive + false\ negative}$$

Abbildung 2: Kennzahlen zur Validierung der UGC Klassifikation [25]

Als Ergebnis der Validierung des illustrierenden Beispiels wurde eine Accuracy von 92.29%, eine Precision von 97.18% und Recall von 89.61% bei der UGC Klassifizierung erreicht. Die Anwendung des angelernten Klassifizierers auf die gesammelte Dokumentenkollektion und Bereinigung von non-UGC führt letztlich zu einem Korpus von 33.866 UGC Dokumenten.

3.4 Extraktion von Markenassoziationen aus Freitexten

Zur Konstruktion des Markenassoziationsnetzwerkes sind zunächst die Netzwerkknoten, also die Markenassoziationen zu bestimmen. Bei der vorliegenden Datengrundlage aus textbasierten UGC Dokumenten, manifestieren sich verbal explizierte Markenassoziationen als Textbestandteile in Form von Worten und Phrasen. Um solche Worte und Phrasen in textbasiertem UGC zu finden, die potenziell Markenassoziationen repräsentieren, ist jedoch die Kenntnis darüber notwendig, wie Verbraucher Markenassoziationen verbal explizieren. Nach Lawson [27] werden Markenassoziationen üblicherweise durch Konzepte repräsentiert, die grundsätzliche Eigenschaften, ähnliche Produkte und Themen, Produktverwendungen oder zusammenfassende Bewertungen beschreiben. Nach unserer Kenntnis finden sich jedoch weder in der Opinion Mining noch der Markenführungsliteratur derzeit Beiträge, welche sich der linguistischen Beschreibung von verbalsprachlich explizierten Markenassoziationen widmen. Dies wäre jedoch für die Konstruktion von Text Mining Verfahren zur Aufdeckung von Markenassoziationen von entscheidendem Wert. Als erste Annäherung wird daher im Rahmen dieses Beitrags folgendes induktives Vorgehen vorgeschlagen, um Merkmale von Markenassoziationen zu beschreiben und diese Beschreibungen für die Konstruktion von Text Mining Verfahren nutzbar zu machen. Auf Grundlage einer Studie, die mit einem traditionell zur Erhebung von Markenassoziationen eingesetzten Verfahren durchgeführt wurde, werden sprachliche Regelmäßigkeiten exploriert, die dann Grundlage des zu konstruierenden Text Mining Verfahrens bilden. Dieses induktive Vorgehen erlaubt es zum einen, aus Bekanntem auf Unbekanntes zu schließen und zum anderen, erste Indikationen der Übereinstimmungsvalidität zu messen, indem die Ergebnisse eines etablierten Erhebungsinstrumentes mit dem hier vorgeschlagenen Erhebungsinstrument verglichen werden [28]. Als traditionelles Erhebungsinstrument wird hierzu die Methode der freien Antwort [29] verwendet, welche zur Verfahrensklasse der assoziativen Verfahren [12] zählt. Bei der freien Antwort werden Probanden nach ihren initialen Gedanken bezüglich eines verbalen Stimulus befragt. Im Rahmen der hier durchgeführten Studie wurden dazu 50 Probanden über die Amazon Mechanical Turk Plattform rekrutiert. Die Probanden stammten aus Nordamerika waren zu 64,7% männlich, im Alter zwischen 25 und 40 Jahren, besaßen zur einen Hälfte einen High School oder Collage Abschluss, zur anderen Hälfte einen Bachelor oder Masterabschluss. In der über ein Online-Formular zugänglichen Befragung wurden die Probanden gebeten, 10 Dinge verbal zu explizieren, die ihnen spontan einfallen wenn sie an „autonomes Fahren“ und „Google“ denken.

Das Ergebnis der quantitativen Vermessung der erhobenen Assoziationen zeigt, dass 65% (329) der insgesamt 500 erhobenen Assoziationen aus einem bis drei Worten bestehen. 42,6% (213) der Assoziationen bestehen gar aus nur einem-, bzw. 15,4% (77)

aus zwei Worten. Daraus lässt sich schließen, dass Verbraucher ihre Assoziationen offenbar vorwiegend als kurze Informationseinheiten wiedergeben. Um nun verbalsprachliche Muster zum Einsatz für das hier vorgeschlagene Text Mining Verfahren feststellen zu können, werden die erhobenen Assoziationen jeweils in eine Sequenz von Wortarten abgebildet. Hierzu wird ein sogenannter Part-of-Speech Tagger (PoS-Tagger) eingesetzt, der jedem Wort innerhalb der Assoziation ein Part-of-Speech Tag (PoS-Tag) -also eine Wortart zuordnet. Im Rahmen dieser Untersuchung wurde der PoS-Tagger von Schmidt [30] auf die einzelnen Assoziationen angewendet und somit für jede Assoziation eine Repräsentation in Form einer PoS-Tag Sequenz erzeugt. Diese Pos-Tag Sequenzen erlauben es nun, die für Assoziationen üblicherweise verwendeten Wortarten, sowie Gemeinsamkeiten in Wortartenabfolgen zu beschreiben. Im Ergebnis zeigt sich, dass nur 15 PoS-Tag Sequenzen mehr als einmal vorkommen, allerdings in der Lage sind, 61,8% der Assoziationen zu konstruieren. Die durchschnittliche Länge dieser Sequenzen beträgt 2,3 Worte. Die verbleibenden 38,2% der Assoziationen werden hingegen von 189 PoS-Tag Sequenzen konstruiert, die jeweils in dieser Form nur einmal vorkommen. Bei Betrachtung der häufigsten Sequenzen zeigt sich, das Assoziationen am meisten durch einzelne und zusammengesetzte Nomen (z.B. <technology>, <safety concerns>) konstruiert werden (~30%), gefolgt von einzelnen Adjektiven (~16%, z.B. <efficient>), sowie der Abfolge von <Adjektiv, Nomen> (6%, z.B. <innovative technology>) oder <Nomen, Adposition, Nomen> (2%, z.B. <safety of driver>. Zusammenfassend kann damit geschlossen werden, dass Nomen und Adjektive die wichtigsten Wortarten zur Bildung von Assoziationen darzustellen scheinen. Die Kombination von Adjektiv und Nomen als beschreibende Attribute wie z.B. „innovative technology“ stellt das am häufigsten vorkommende Muster bei Assoziationen dar. Diese Erkenntnisse werden nun genutzt um ein Text Mining Verfahren zu konstruieren. Hierzu wird ähnlich zur oben durchgeführten Charakterisierung von Markenassoziationen ein Vorgehen aus Tokenisierung, Part-of-Speech-Tag (PoS-tag) Sequenzuntersuchung, sowie Sentimentanalyse vorgeschlagen. Zunächst werden dazu die UGC Dokumente in kleinere Bestandteile mittels Tokenisierung [31] segmentiert um für jedes Dokument eine Menge von Sätzen, sowie für jeden Satz eine Menge von Worten erhalten. Da nicht sichergestellt werden kann, dass in den gesammelten UGC Dokumenten Verbraucher ausschließlich über Google und autonomes Fahren sprechen und alle extrahierten Assoziationen sich auf diesen Untersuchungsgegenstand beziehen, wird die Menge der mittels Tokenisierung ermittelten Sätze auf jene reduziert, in denen sich Hinweise auf den Untersuchungsgegenstand finden (z.B. die Nennung von Google). Auf Grundlage dieser selektierten Sätze, wird dann eine PoS-tag-Sequenzuntersuchung zur Erhebung der Markenassoziationskandidaten durchgeführt. Es wird dazu jedem einzelnen Wort eines Satzes ein PoS-Tag mittels des Ansatzes von Schmidt [30] zugewiesen. Dies erlaubt es, für jeden aus Worten bestehenden Satz eine Sekundärrepräsentation zu erstellen, welche allein aus einer Sequenz von PoS-tags besteht. Diese PoS-tag Sequenzen erlauben es nun, Muster zur Extraktion von Markenassoziationen anzuwenden. Auf Grundlage der oben beschriebenen sprachlichen Charakteristika von Markenassoziationen wird hierbei beispielsweise das PoS-tag Muster <Adjektiv, Nomen> verwendet um Wortsequenzen wie <technological innovation> als Markenassoziationskandidaten zu extrahieren und in einer Datenbank zu speichern.

3.5 Sentimentanalyse

Bei Markenassoziationen ist die emotionale Orientierung (z.B. positiv, negativ, neutral) von besonderem Interesse. Keller [1] beispielsweise beschreibt das kundenorientierter Markenwert nicht nur durch die im Gedächtnis der Konsumenten vorhandenen starken und einzigartigen, sondern auch vorteilhaften – also positiven - Markenassoziationen entsteht. Daher sollte für die extrahierten Markenassoziationskandidaten die Valenz, also das sog. Sentiment bestimmt werden. Liu [32] liefert hierzu einen umfassenden Überblick über verschiedene Techniken. Da das Forschungsfeld der aspektorientierten Sentimentanalyse hinsichtlich der Erkennungsgenauigkeit schnell voranschreitet, sieht die hier vorgeschlagene Methodik das Einsetzen jeweils aktueller Verfahrensweisen der aspektorientierten Sentimentanalyse vor. Im Rahmen der illustrierenden Studie wird ein einfacher lexikonbasierter Ansatz vorgeschlagen und genutzt, welcher für jede extrahierte Adjektiv-Nomen Kombination ein referenzierendes Lexikon wie z.B. SenticNet [33] zur Bestimmung der emotionalen Orientierung anhand des Adjektivs verwendet. Als Ergebnis der vorangegangenen Schritte wird nun für jedes Dokument eine Menge von Markenassoziationskandidaten (z.B. „leading technology“, „innovative brand“ oder „threatened privacy“) erhalten. Das jeweilige Bezugswort (z.B. das Nomen: „technology“, „brand“, „privacy“) repräsentiert dabei die Markenassoziation, die referenzierenden Worte (z.B. Adjektiv: „leading“, „innovative“, „threatened“) werden als tonalitätstreibende Aspekte verwendet.

3.6 Bestimmung der Assoziationsstärke

Die Assoziationsstärke beeinflusst wie schnell und einfach Assoziationen aus dem semantischen Gedächtnis eines Kunden abgerufen werden können [1]. Starke Markenassoziationen, sogenannte „Branded Features“ [10], sind also jene, welche die meisten Kunden mit der Marke in Verbindung bringen und an die sie als erstes im Markenzusammenhang denken. Sie definieren die Marke aus Kundensicht und werden daher auch als „Essenz“ der Marke bezeichnet [34]. Aus diesem Grunde stellt die Stärke von Assoziationen ein wichtiges Attribut zur Beurteilung von Markenassoziationsnetzwerken dar [1]. Angelehnt an traditionelle, qualitative Erhebungsmethoden [34], bei denen die Stärke einer Assoziation anhand der Häufigkeit der Nennung einer Assoziation durch Studienteilnehmer ermittelt wird, soll im Rahmen der hier vorgeschlagenen Methodik die Stärke einer Assoziation anhand ihrer Vorkommenshäufigkeit charakterisiert werden. Hierzu wird ein 2-Mode Netzwerk konstruiert, welches zur Bestimmung der Assoziationsstärke mittels der Gradzentralität [35] dienen soll. Zur Konstruktion des Netzwerkes werden zwei Typen von Knoten unterschieden. Dokumentknoten existieren für jedes UGC Dokument. Als zweiter Typ existiert für jede Markenassoziation die aus einem Dokument extrahiert wurde ein Assoziationsknoten. Jeder Assoziationsknoten eines Dokumentes wird mit jeweils dem korrespondierenden Dokumentknoten über eine Dokument-zu-Assoziation Kante verbunden. Wird eine Assoziation (welche über ihren Namen sowie die emotionale Orientierung zu identifizieren ist) in mehreren unterschiedlichen Dokumenten genannt, so existiert der Assoziationsknoten dennoch nur einmal, wird jedoch mit jedem Dokumentknoten (und somit dem korrespondierenden Dokument) mit einer Dokument-zu-Assoziationskante verbunden. Abbildung 3 illustriert die Erstellung dieses 2-mode Netzwerk. Hierbei entspricht die Gradzentralität des Assoziationsknoten [35, 36] der Anzahl der Dokumente in denen die jeweilige Assoziation existiert. Um Vergleichbarkeit herzustellen, wird die Gradzentralität mit der Gesamtanzahl von Dokumenten normalisiert.

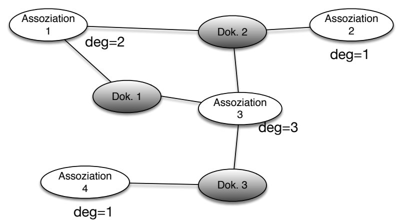


Abbildung 3: 2-mode Netzwerk zur Bestimmung der Assoziationsstärke

3.7 Bestimmung der Verbindung zwischen Assoziationen

Während im vorangegangenen Abschnitt das vorgeschlagene Verfahren zur Ermittlung der Stärke von Assoziationen dargestellt wurde, wird nun das Vorgehen erläutert, um Verbindungen zwischen Assoziationen sowie deren Stärke darzustellen. Aus kognitionspsychologischer Sicht wird die Verbindungsstärke zwischen Assoziationen charakterisiert durch die Stärke mit der diese miteinander im Gedächtnis verknüpft werden. Sie beeinflusst damit Aktivierungsausbreitung („spreading activation") [1], also die Wahrscheinlichkeit das weitere Assoziationen im Gedächtnis des Verbrauchers aktiviert werden, wenn eine einzelne Assoziation z.B. durch Marketing-Mix Aktivitäten aktiviert wird [4, 37]. Im Gegensatz zu traditionellen Consumer-mapping Verfahren [17] können die Verbindungen zwischen Assoziationen im hier vorgeschlagenen Verfahren nicht direkt durch Befragung von Probanden ermittelt werden. Daher wird angelehnt an das von Teichert und Schöntag [38] vorgestellte Verfahren, die Annahme getroffen, dass aus der implizit gemeinsamen Nennung von Assoziationen auf vorhandene Assoziationsnetzwerke geschlossen werden kann. Pro Dokument wird dazu jeweils ein vollständiger Graph aus den dort extrahierten Assoziationen konstruiert (Abbildung 4). Dem unterliegt die Annahme, dass die vom Verbraucher innerhalb eines UGC Dokumentes verbal explizierten Assoziationen im Gedächtnis in einem gemeinsamen Bezug stehen. Über die gesamte Dokumentenkollektion wird dann gezählt, wie oft jeweils ein Paar von gleichen Assoziationen in unterschiedlichen Dokumenten gemeinsam auftritt. Jeweils zwischen diesen Assoziationspaaren wird dann eine gewichtete Kante erzeugt, deren Gewicht dieser Kookkurrenz entspricht – also der Anzahl der Dokumente in denen die beiden Assoziationen gemeinsam auftreten. Da häufig vorkommende Assoziationen in dieser Betrachtung eine höhere Wahrscheinlichkeit haben mit anderen weniger häufig vorkommenden Assoziationen aufzutreten, kann ein Korrekturfaktor einbezogen werden, um solche Verbindungen hervorzuheben die häufiger als zu erwarten wären auftreten.

Statistische Unabhängigkeit vorausgesetzt, entspricht die Wahrscheinlichkeit, dass zwei Assoziationen zufällig gemeinsam auftreten, dem Produkt der relativen Häufigkeit (bzw. normalisierten Gradzentralität) der beteiligten Assoziationen. Diese zu erwartende Wahrscheinlichkeit kann von der relativen Häufigkeit des gemeinsamen Auftretens (Quotient aus absoluter Kookkurrenz zweier Assoziationen und der Vorkommenshäufigkeit der schwächeren Assoziation) abgezogen werden. Abbildung 4 zeigt das resultierende Assoziationsnetzwerk auf Grundlage von zwei einzelner vollständiger Dokumentnetzwerke.

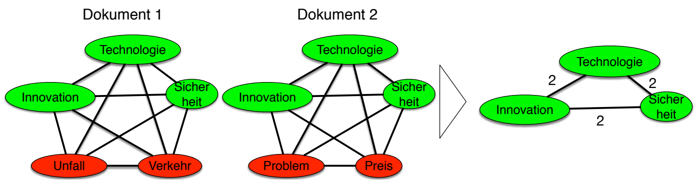


Abbildung 4: Per Dokument und Kookkurenznetzwerke

3.8 Konstruktion des Markenassoziationsnetzwerkes

Zur Konstruktion des Markenassoziationsnetzwerkes werden die Netzwerke aus den Abschnitten 3.6 und 3.7 kombiniert. Hierbei wird das Netzwerk aus Abschnitt 3.6 zum einen dazu verwendet, mittels der normalisierten Gradzentralität die Stärke der Assoziationsknoten zu beschreiben und zum anderen um einen Schwellwert zur späteren Berücksichtigung der Assoziationen im Markenassoziationsnetzwerk festzulegen. Das Netzwerk aus Abschnitt 3.7, welches die Verbindungen zwischen Assoziationen beschreibt, wird dazu um die Information der Assoziationsstärke ergänzt.

Im Bezug zur illustrierenden Studie zeigt Abbildung 5 das Netzwerk der Top-50 Assoziationen für Google im Kontext des autonomen Fahrens. Als Schwellenwert für das gemeinsame Auftreten zweier Assoziationen wurde 18 gewählt. Die Assoziationsknoten wurden gemäß des Anteils positiver (grün), negativer (rot) sowie neutraler (orange) Tonalität eingefärbt. Die Gradzentralität der Assoziationen ist im Zentrum der Assoziationsknoten dargestellt. Insgesamt kann festgestellt werden, dass die im Markenassoziationsnetzwerk in Abbildung 5 vorhandenen Assoziationen inhaltlich ebenfalls bei der befragungsbasierten Studie aus Abschnitt 3.4 wiedergefunden werden können, was eine erste schwache Indikation der Vergleichsvalidität sein könnte. Allgemein sind im Netzwerk in Abbildung 5 zwei schwach miteinander vernetzte Subnetzwerke zu beobachten. Das erste Subnetzwerk beinhaltet Assoziationen die dem Themenfeld des „autonomen Fahrens" zugeordnet werden können. Diese Assoziationen sind entweder direkt oder in zweiter Ordnung mit der Assoziation „car" verknüpft. Direkt verknüpfte Assoziationen lassen dabei auf die von Google durchgeführten Prototypentests schließen wie z.B. „testing", „prototype", „area" und „state". Andere Assoziationen betonen gesellschaftliche, soziale oder persönliche Aspekte. Hierzu zählt beispielsweise die Assoziation „safe", die stark mit der Assoziation „car", sowie „driver" verknüpft ist. Das zweite Subnetzwerk hingegen zeigt Assoziationen, die nicht unbedingt auf das autonome Fahren schließen lassen, jedoch scheinbar starke Markenattribute darstellen, die von Verbrauchern in kontextunabhängig wiedergegeben werden. Hierzu zählen „computer“, „company“, „services“, „search“, „computer“, „project“, „technology“, „wearables“, „Google Glass“, „devices“ oder „android“. Auffällig ist, dass das erste, dem autonomen Fahren nahe Subnetzwerk, schwach über die Verbindung der zentralen Assoziation „car“ mit der Assoziation „project“ des zweiten Subnetzwerkes verbunden ist, autonomes Fahren also möglicherweise als eines von mehreren Projekten Googles wahrgenommen wird.

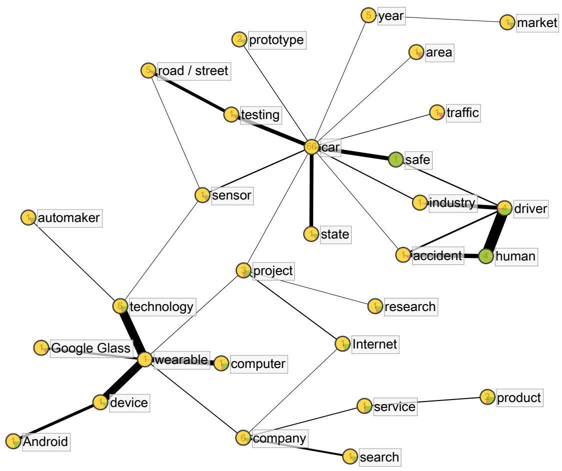


Abbildung 5: Assoziationsnetzwerk Google im Kontext autonomes Fahren

4 Einsatz zur Imagediagnostik

Die Verwendung netzwerkanalytischer Verfahren auf Markenassoziationsnetzwerken und die Interpretation in Bezug auf Markenführungsmaßnahmen wurde insbesondere von Henderson auf Grundlage von durch Befragung erhobenen Markenassoziationsnetzwerken untersucht [10, 11]. Die im Rahmen dieses Beitrags vorgestellte Methodik zur Erhebung von Markenassoziationsnetzwerken aus nutzergenerierten Inhalten können ebenfalls als Grundlage solcher Untersuchungen dienen. Henderson et al. [10] unterscheiden grundsätzlich zwischen Intra-Netzwerkanalysen und Inter-Netzwerkanalysen [10]. In der Intra-Netzwerkanalyse wird auf einzelne Assoziationsknoten fokussiert. Die Gradzentralität wurde hier beispielsweise verwendet um besonders starke Assoziationen sog. „Branded Features“ zu untersuchen. Da diese Branded Features von Verbrauchern als besonders wichtige Markenattribute wahrgenommen werden, sollten diese bei Markenführungsmaßnahmen besonders herausgestellt werden. Henderson et al. [10] verwenden hier beispielsweise die Maße der Gradzentralität (Anzahl der Kanten zu anderen Assoziationsknoten) sowie der Betweenness-Zentralität (Anzahl der kürzesten Wege, die über den Assoziationsknoten zu anderen Assoziationen führen) um die wichtigsten Branded Features zu beschreiben. Im Bezug auf die hier angeführte illustrierende Studie könnte die Gradzentralität Hinweise darauf geben, welche Assoziationen von Verbrauchern mit

Google im Bezug auf das autonome Fahren thematisiert werden und somit wichtige Markenattribute darstellen. Die Verwendung der Betweenness Zentralität könnte verwendet werden, um zu untersuchen wie gem. der Idee des Spreading Activation Networks [4] die Wahrscheinlichkeit besteht, dass eine gewisse Assoziation aktiviert wird, wenn eine andere Assoziation im Markenassoziationsnetzwerk aktiv ist. Marketingmaßnahmen können diese Assoziation dann ansprechen (oder nicht ansprechen) um dafür zu sorgen (oder zu verhindern), dass andere Assoziationen aktiv werden. Neben der Intra-Netzwerkanalyse bietet sich ferner die Möglichkeit unterschiedliche Netzwerke, oder Segmente von Netzwerken zu vergleichen um unterschiedliche Marktsegmente gegenüberzustellen oder Gemeinsamkeiten und Unterschiede in der Verbraucherwahrnehmung herauszuarbeiten. Mit einem geeigneten Informationssystem, welches das hier vorgestellte Verfahren implementiert, könnten beispielsweise Netzwerke anhand der Beiträge einer bestimmten Webseite oder eines Zeitraums segmentiert werden, da für nahezu jedes Dokument die Webseite sowie der Zeitpunkt der Veröffentlichung bekannt ist. [10] führen als Beispiel für Inter-Netzwerkanalysen auch das Maß der Netzwerkdichte als Vergleichskriterium an. Ein sehr dichtes Netzwerk für eine einzelne Marke könnte hierbei eine Indikation auf eine unklare Positionierung geben und somit auf eine mögliche Verwässerung des Markenwertes hinweisen [39]. Gleiches kann angenommen werden, wenn ein Markenassoziationsnetzwerk im Zeitverlauf an Dichte gewinnt [11].

5 Diskussion und Ausblick

Im Rahmen dieses Beitrages wurde ein Verfahren vorgeschlagen, welches potenzielle Markenassoziationen aus UGC extrahiert und somit zur Methodenentwicklung im Bereich der automatisierten Verbraucherforschung beiträgt. Das Verfahren wird hierbei aus der Literatur zur Markenführung und Kognitionspsychologie abgeleitet, ist im Vergleich zu befragender Forschung aufgrund Automatisierung ggf. schneller und potenziell kostengünstiger [13]. Sofern den UGC Beiträgen Metadaten wie Zeitstempel angefügt sind, ist eine Analyse mit dem Verfahren auch rückblickend durchführbar. Den Hauptnutzen besitzt das Verfahren jedoch darin, dass es UGC als eine bisher nicht oder nur schwer holistisch zugreifbare Datenquelle für die Markt- und Verbraucherforschung zugänglich macht. Es lässt sich festhalten, dass mittels unterschiedlicher Intra- sowie Inter-Netzwerkanalysen potenziell Marketing-relevante Informationen abgeleitet werden könnten. Dem vorgelagert stellt sich jedoch zunächst die wichtige Frage welche Gegenstand der folgenden Forschung sein sollte, ob die hier vorgestellten Netzwerke validiert werden können. Ein Ansatz wäre es, die Übereinstimmungsvalidität zu überprüfen: Hierbei wären mit einem „etabliert und als valide geltenden" Verfahren Assoziationsnetzwerke zu erheben und diese mit den Ergebnissen des hier vorgestellten Verfahrens zu vergleichen. Erste Indikationen wurden hier im Vergleich der erhobenen Assoziationen zu einer befragungsorientierten Studie geliefert. Weiterhin besteht die Frage, wie geeignete Schwellwerte für das Berücksichtigen von erhobenen Assoziationen (z.B. anhand der Gradzentralität) zu setzten sind, da im Vergleich zu befragenden Instrumenten hier nicht 10 oder 20 Assoziationen sowie Relationen zwischen Assoziationen erhoben werden, sondern Hundertausende bis zu mehreren Millionen. Letztlich stellt das hier vorgestellte Verfahren einen ersten Ansatz dar, Markenassoziationsnetzwerke aus der digitalen Flut von nutzergenerierten Inhalten als diagnostisches Instrument für die Markenführung verfügbar zu machen. Konkrete Instanziierungen dieses Verfahrens in Form von Informationssystemen (z.B. zum Marketing-Controlling) würden dem unternehmensinternen Methodenkanon zur Marktforschung und Markenführung eine neue Datenquelle erschließen und somit neue, substantielle Erkenntnismöglichkeiten über die Markenwahrnehmung der Verbraucher im Internet eröffnen. Wir hoffen daher, dass andere diesen Ansatz aufgreifen und weitere Forschung im Bereich der textanalysebasierten Erforschung von Markenassoziationsnetzwerken stattfindet.

References

1. Keller, K.: Conceptualizing, measuring, and managing customer-based brand equity. Journal of Marketing. 57, 1–22 (1993).
2. Esch, F.-R., Herrmann, A., Sattler, H.: Marketing-Eine managementorientierte Einführung. Auflage, Vahlen Vl. (2008).
3. Esch, F.-R., Langner, T., Rempel, J.E.: Ansätze zur Erfassung und Entwicklung der Markenidentität. In: link.springer.com. pp. 103–129. Gabler Verlag, Wiesbaden (2005).
4. Collins, A.M., Loftus, E.F.: A spreading-activation theory of semantic processing. Psychological review. 82, 407 (1975).
5. Aaker, D.A.: Managing Brand Equity: Capitalizing on the Value of a Brand Name. New York. (1991).
6. Keller, K.L.: Brand synthesis: The multidimensionality of brand knowledge. Journal of consumer research. 29, 595–600 (2003).
7. Keller, K.L., Lehmann, D.R.: Brands and Branding: Research Findings and Future Priorities. Marketing Science. 25, 740–759 (2006).
8. Anderson, J.R., Bower, G.H.: Human Associative Memory. Lawrence Erlbaum (1980).
9. Krishnan, H.S.: Characteristics of memory associations: A consumer-based brand equity perspective. International Journal of Research in Marketing. 13, 389–405 (1996).
10. Henderson, G., Iacobucci, D., Calder, B.J.: Brand diagnostics: Mapping branding effects using consumer associative networks. European Journal of Operational Research 111. 306–327 (1998).
11. Henderson, G.R., Iacobucci, D., Calder, B.J.: Using network analysis to understand brands. Advances in Consumer Research. 29, 397–405 (2002).
12. Farsky, M.: Methoden zur Messung des Markenimages: State of the Art. Research Papers on Marketing and Retailing, Hamburg (2007).
13. Urban, G.L., Hauser, J.R.: "Listening in" to find and explore new combinations of customer needs. Journal of Marketing. 68, 72–87 (2004).
14. Egger, M., Lang, A.: A Brief Tutorial on How to Extract Information from User-Generated Content (UGC). Künstl Intell. 27, 53–60 (2013).
15. Decker, R., Trusov, M.: Estimating aggregate consumer preferences from online product reviews. International Journal of Research in Marketing. 27, 293–307 (2010).
16. Netzer, O., Feldman, R., Goldenberg, J., Fresko, M.: Mine Your Own Business: Market-Structure Surveillance Through Text Mining. Marketing Science. 31, 521–543 (2012).
17. John, D., Loken, B., Kim, K., Monga, AB: Brand concept maps: a methodology for identifying brand association networks. Journal of Marketing. 43, 549–563 (2006).
18. Schnittka, O., Sattler, H., Zenker, S.: Advanced brand concept maps: A new approach for

evaluating the favorability of brand association networks. International Journal of Research in Marketing. 29, 265–274 (2012).
19. Brandt, C., de Mortanges, C.P., Bluemelhuber, C., van Riel, A.C.: Associative networks: A new approach to market segmentation. Int. J. Market Res. 53, 187–207 (2011).
20. Crawford Camiciottoli, B., Ranfagni, S., Guercini, S.: Exploring brand associations: an innovative methodological approach. European Journal of Marketing. 48, 1092–1112 (2014).
21. Liu, B., Hu, M., (null): Opinion Observer: Analysing and Comparing Opinions on the Web. Presented at the Proceedings of the 14th international conference on World Wide Web, Chiba, Japan April 1 (2005).
22. Popescu, A., Etzioni, O.: Extracting product features and opinions from reviews. Presented at the Proceedings of the Conference on Human Language Technology and Empirical Methods in Language Processing (2005).
23. Gensler, S., Völckner, F., Egger, M., Fischbach, K., Schoder, D.: Listen to Your Customers: Insights into Brand Image Using Online Consumer-Generated Product Reviews. International Journal of Electronic Commerce. 20, 112–141 (2015).
24. Kohlschütter, C., Fankhauser, P., Nejdl, W.: Boilerplate detection using shallow text features. Presented at the Proceedings of the third ACM international conference on Web search and data mining (2010).
25. Egger, M., Lang, A., Schoder, D.: Who Are We Listening to? Detecting User-generated Content (UGC) on the Web. Presented at the ECIS 2015 Completed Research Papers, Münster August 26 (2015).
26. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. 20, 273–297 (1995).
27. Lawson, R.: Consumer Knowledge Structures: Networks and Frames. NA - Advances in Consumer Research Volume 25. (1998).
28. Association, A.P.: Technical recommendations for psychological tests and diagnostic techniques. Part 2 Psychological Bulletin (1954).
29. Olson, J.C., Muderrisoglu, A.: The Stability of Responses Obtained By Free Elicitation: Implications For Measuring Attribute Salience and Memory Structure. Advances in Consumer Research. 6, 269–275 (1979).
30. Schmid, H.: Probabilistic part-of-speech tagging using decision trees. Presented at the Proceedings of the international conference on new methods in language processing (1994).
31. Feldman, R., Sanger, J.: The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge University Press (2006).
32. Liu, B.: Sentiment Analysis and Opinion Mining. Computational Linguistics. 1–5 (2014).
33. Cambria, E., Olsher, D., Rajagopal, D.: SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis. Presented at the Twenty-eighth AAAI conference on artificial intelligence (2014).
34. Van Rekom, J., Jacobs, G.: Measuring and managing the essence of a brand personality. Marketing Letters. (2006).
35. Wasserman, S., Faust, K.: Social network analysis: Methods and applications. Cambridge University Press (1994).
36. Scott, J.: Social network analysis. Sage (2012).
37. Anderson, J.R.: A spreading activation theory of memory. Journal of verbal learning and verbal behavior. 22, 261–295 (1983).
38. Teichert, T.A., Schöntag, K.: Exploring consumer knowledge structures using associative network analysis. Psychol. Mark. 27, 369–398 (2010).
39. Barbara Loken, D.R.J.: Diluting Brand Beliefs: When Do Brand Extensions Have a Negative Impact? Journal of Marketing. 57, 71–84 (1993).

Leveraging RFID Data Analytics for the Design of an Automated Checkout System

Matthias Hauser[1], Sebastian Günther[1], Christoph M. Flath[1], and Frédéric Thiesse[1]

[1] University of Würzburg, Department of Business Management, Würzburg, Germany
matthias.hauser@uni-wuerzburg.de, sebastian.guenther@aol.de
christoph.flath@uni-wuerzburg.de, frederic.thiesse@uni-wuerzburg.de

Abstract. Traditional checkout systems are labor-intensive and can be a source of great frustration for customers having to wait in line. In contrast, automated checkout systems seamlessly scan, total and charge a customer's purchase to a registered payment account while they are simply leaving the store. We focus on the main challenge of automatically detecting customer purchases. To this end, we develop a checkout system that leverages data mining techniques to (i) reliably and timely detect items leaving the shopping floor area and (ii) assign them to individual customers. We demonstrate the system's feasibility using a large data set collected in the laboratory under real-world conditions.

Keywords: Data Analytics, RFID, Internet of Things, Automated Checkout

1 Introduction

While RFID solutions are commonly used in retail supply chains for automatic detection of logistical units in upstream and backroom processes, front store customer-facing applications have not progressed beyond pilot trials. RFID solutions for this "last mile" of the supply chain are especially challenging because of the sheer number and variety of simultaneously moving objects. Complexity is further increased by the manner in which objects are transported (stacked, in bags, etc.), unpredictable customer behavior, and suboptimal store layouts. Against this backdrop, we consider RFID-based automated checkout systems, which promise great value potential for the retail industry [1]. Our research objective is to decompose the challenge of such systems into tractable sub problems and demonstrate the feasibility of a pilot implementation.

2 Automated Checkout System Requirements

Traditional checkout systems are labor-intensive and can be a great source of frustration for customers when having to wait in line. To reduce costs retailers have started adopting self-service systems which enable shoppers to scan, bag and pay for their purchases with little or no help from store personnel [2, 3]. These systems, however, offer hardly any improvements over traditional checkout with respect to the customer

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Hauser, M.; Günther, S.; Flath, C.; Thiesse, F. (2017): Leveraging RFID Data Analytics for the Design of an Automated Checkout System, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1201-1204

experience, potentially offering an even worse experience. In contrast, automated checkout systems scan, total and charge a customer’s purchases to a registered payments account while they are simply leaving the store.

In our research, we focus on the main challenge of automatically detecting customer purchases. Here, we must cope with two problems: The system (i) has to reliably detect all purchased products and (ii) assign these to individual customers. Undetected products cause direct losses to the retailer and inventory inaccuracy resulting in out-of-stock situations [4]. Incorrectly assigning items to a customer, on the other hand, may lead to customer dissatisfaction and interruptions of in-store operations [5].

Academic literature on automated checkout systems is sparse. To the best of our knowledge, only one group of researchers have developed a system (“MyGrocer”) that addresses the previously mentioned challenges [6, 7]. This solution relies on shopping carts equipped with RFID readers that detect objects placed in the carts. Similar to this approach, we also rely on RFID for the detection of products. In contrast to optical barcodes, RFID allows identification of products at the item level. In addition, identification does not require a direct line of sight between the tag and the reader device, which allows for simultaneous bulk detection of multiple objects. In contrast to the previous studies, we decided against RFID-equipped carts for several reasons: First, shopping carts are not used in all types of retail stores (e.g., fashion retail stores) which limits the generalizability. Second, the costs for equipping shopping cart with RFID readers are very high [7]. In addition, there are high operating costs for regular service of the carts. Instead, we propose an automated checkout system that is implemented at the exits of retail stores and does not require customers to use specific devices. This system facilitates the detection of items in carts and items carried by customers.

The MyGrocer carts only need to detect items within them. In this case, sufficient data quality can be achieved by selecting RFID antennas with read ranges of only 20cm. In our case, the antennas need to detect items that leave a store through an exit gate, which requires antennas with large read range and high power. Unfortunately, this in turn leads to the detection of RFID tags carried nearby by the gate instead of through the gate. Secondly, assigning items to customers can hardly be realized with hardware-based approaches. Customers would have to wait in line and pass the gate one after the other to avoid purchases of different customers being read at the same time.

3 Description of the Automated Checkout Artifact

We pursue a design-oriented research approach to create an automated checkout artifact that (i) reliably and timely detects items leaving a store and (ii) correctly assigns them to individual customer. The architecture combines hardware components and software components. The hardware component consists of two RFID reader installations, a ceiling-mounted system which tracks items in the store and a gate-mounted system that helps to detect items that are leaving the store (see Figure 1). This infrastructure collects low-level RFID data that is then processed by the software components.

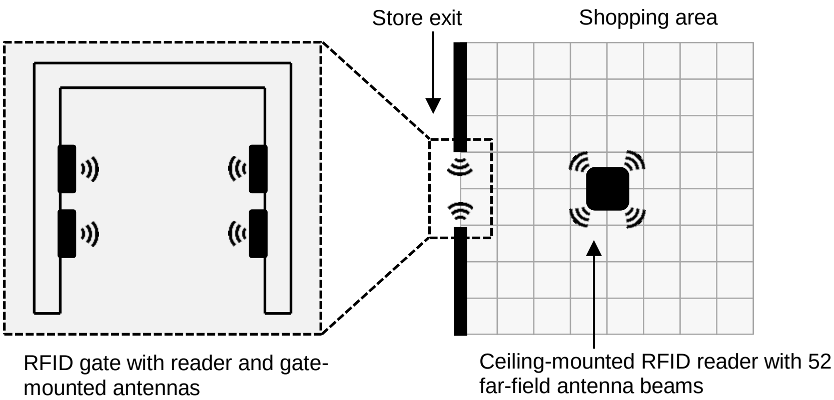


Figure 1. Hardware infrastructure with two distinct RFID reader installations

There are two distinct software functionalities. The *first* software component uses data mining techniques to reliably and timely detect items that are leaving the store. To this end, we extract features from the low-level RFID data, which contain information regarding observed real world events. One simple example is the maximum received signal strength value measured in a series of detections of a particular tag within a certain time interval. With these features we train classification models that are able to distinguish between items that leave the store through the exit gate and others. We follow the conceptual approach described in Hauser et al. [8]. However, we go beyond the state-of-the-art by considering run fragments in real-time instead of evaluating completed paths ex-post. The *second* software component assigns items leaving the store (identified by the first component) to individual customers. To this end, we infer item paths in the shopping area and then apply cluster analysis to group them. The procedure rests on the assumption that the paths of items purchased by one customer are more similar to each other than to paths of other items.

4 Expected Contribution and Future Work

We use a large data set collected in a laboratory environment for instantiation and evaluation of the artifact. Our experimental setup takes into account the limited process control in retail stores by considering multiple walking paths, varying numbers of persons and items as well as different movement speeds. Our experimental design includes 18 different tests. We repeated each of the tests 50 times. We then perform 5-fold validation to ensure representative results. We found that the artifact reliably and timely detects all shopping baskets carried out of the shopping floor area. On average, the baskets were detected 0.41 seconds before customers stepped out of the sales floor area. In addition, 94% of the shopping baskets assigned by the artifact contained the correct items.

These misclassifications arise in two particularly challenging test scenarios where multiple customers approach the exit gate simultaneously on very similar (almost identical) movement paths. In practice such a situation could easily arise when friends are shopping together which highlights the limitations of the pilot implementation. Nonetheless, our initial study demonstrates the fundamental feasibility of RFID-based automated checkout. Going forward, we want to evaluate our artifact in different test environments to evaluate its generalizability. In addition, we want to include additional, richer test settings (e.g., scenarios in which customers take objects from shelves that are placed near the exits). Our ultimate objective is to ensure feasibility under real-world conditions.

5 Acknowledgement

This research was carried out under the Research Project SERAMIS (Sensor-Enabled Real-World Awareness for Management Information Systems), funded by the EU under the VII Framework Research Program.

References

1. Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., Aharon, D.: The Internet of Things: Mapping the Value Beyond the Hype. McKinsey Global Institute (2015).
2. Orel, F.D., Kara, A.: Supermarket self-checkout service quality, customer satisfaction, and loyalty: Empirical evidence from an emerging market. J. Retail. Consum. Serv. 21, 118–129 (2014).
3. Litfin, T., Wolfram, G.: New Automated Checkout Systems. In: Krafft, M. and Mantrala, M.K. (eds.) Retailing in the 21st Century: Current and Future Trends. pp. 143–157. Springer Berlin Heidelberg, Berlin, Heidelberg (2006).
4. Kang, Y., Gershwin, S.B.: Information inaccuracy in inventory systems: stock loss and stockout. IIE Trans. 37, 843–859 (2005).
5. Hayes, R., Blackwood, R.: Evaluating the effects of EAS on product sales and loss: Results of a large-scale field experiment. Secur. J. 19, 262–276 (2006).
6. Kourouthanassis, P., Roussos, G.: Developing Consumer-Friendly Pervasive Retail Systems. IEEE Pervasive Comput. 2, 32–39 (2003).
7. Roussos, G., Kourouthanasis, P., Spinellis, D., Gryazin, E., Pryzbliski, M., Kalpogiannis, G., Giaglis, G.: Systems Architecture for Pervasive Retail. In: Proceedings of the 2003 ACM Symposium on Applied Computing. pp. 631–636. ACM, New York, NY, USA (2003).
8. Hauser, M., Zügner, D., Flath, C.M., Thiesse, F.: Pushing the limits of RFID: Empowering RFID-based Electronic Article Surveillance with Data Analytics Techniques. In: Proceedings of the Thirty Sixth International Conference on Information Systems (2015).

Decision Support for the Automotive Industry: Forecasting Residual Values using Artificial Neural Networks

Christoph Gleue[1], Dennis Eilers[1], Hans-Jörg von Mettenheim[1], Michael H. Breitner[1]

[1] Leibniz Universität Hannover, Institut für Wirtschaftsinformatik, Hannover, Germany
{gleue,eilers,mettenheim,breitner}@iwi.uni-hannover.de

Abstract. The leasing business is one of the most important distribution channels for the automotive industry. This implies that forecasting accurate residual values for the vehicles is a major factor for determining monthly leasing rates: Either a systematic overestimation or underestimation of future residual values can incur large potential losses in resale value or, respectively, competitive disadvantages. In this paper, an operative DSS with the purpose of facilitating residual value related management decisions is introduced, with a focus on its forecasting capabilities. Practical implications are discussed, a multi-variate linear model and an artificial neural network approach are benchmarked and further, the effects of price trends and seasonal influences are investigated. The analysis is based on more than 150,000 data sets from a major German car manufacturer. We show that artificial neural network ensembles with only a few input variables are capable of achieving a significant improvement in forecasting accuracy.

Keywords: Decision Support Systems, Business Intelligence, Artificial Neural Networks, Residual Value Forecasts, Car Leasing

1 Introduction and Motivation

The leasing market is an important business segment for car manufacturers. Automotive assets, which include passenger cars and commercial vehicles, accounted for a volume of 65% (178.2 billion Euros) of total new leasing contracts granted in 2014, which makes it the largest individual asset segment of the European leasing market [1]. Almost a third of all new cars sold by German premium brands in 2015 was financed with leasing [2]. Hence, this business model provides tremendous market opportunities, particularly for car manufacturers and leasing companies. Nevertheless, there are also risks which are difficult to quantify. The focus of our research is the so-called residual value risk [3]. While the customer pays a contractually-fixed leasing rate over the entire period of use, the residual value of the vehicle is uncertain until the end of the leasing period. In order to compensate the loss in value of the vehicles with adequate leasing payments, forecasts of residual values must be as exact as possible. Hence, accurate forecasts of future residual values constitute a critical success factor for competitiveness in the leasing market. Either a systematic over-estimation or under-

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Gleue, C.; Eilers, D.; von Mettenheim, H.-J.; Breitner, M. H. (2017): Decision Support for the Automotive Industry: Forecasting Residual Values using Artificial Neural Networks, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1205-1219

estimation of future residual values would have negative consequences. If an overestimate is determined, the leasing payments may be lower, but not sufficient to compensate the loss in value of the vehicle. If predicted residual values are systematically too low, higher leasing rates must be set. This means that either avoidable competitive disadvantages occur or leasing rates must be substantially subsidized by car manufacturers. Car manufacturers nowadays often already have a sufficiently good data history of completed transactions in the leasing sector. They also record resale prices in the used car market for the corresponding vehicles. This potential, however, is often not sufficiently utilized to improve predictions of residual values.

In our research, we introduce an operational Decision Support System (DSS) which was developed and successfully implemented in cooperation with a large German car manufacturer. Although the incorporated higher-level analyses, management reports and visualizations are equally important parts of the aforementioned DSS, they all rely on proper data handling and a reliable forecasting methodology. Thus, in this paper we focus on the forecasting module of the system. The residual value forecast is a typical regression problem in which certain characteristics such as the age and mileage of a particular vehicle determine its residual value. The study is based on more than 150,000 records of completed leasing transactions and resale values over a period of four years (2011-2015). Linear models and Artificial Neural Networks (ANN) are implemented in this study. For this purpose, the development of vehicle values depending on the model age and time factor is of particular interest. The model age measures the time between the market launch of a new model class and the resale date of a specific vehicle belonging to this class. The continuously increasing model age means that the expected residual value of two identical vehicles of exactly the same age and with the same mileage is not constant over time. In addition, prices are influenced by general market conditions, accounted for by a time component in our analysis. Furthermore, the residual values on the used car market are subject to seasonal fluctuations within a year (a critical factor for ANN forecasts [4]). In order to allow for an unbiased forecast of estimated residual values over several years, these influences are examined more closely.

The remainder of this paper is structured as follows. The next section 2 provides an overview of the existing literature in the field of residual value forecasts. The structure of the DSS and the available input data are presented and explained in section 3. Section 4 specifies an exploratory analysis about trends and seasons in the used car market and subsequently introduces the forecasting model based on ANNs. The results of the analysis are presented in section 5. Section 6 provides a discussion and addresses limitations, implications for practical use and further research opportunities. Section 7 concludes with a short summary.

2 Related Work

This section presents the existing literature in the field of residual value forecasts. Due to the exclusivity of the data, only few studies on this subject exist. This fact, however, also highlights the enormous research potential. One of these studies is provided by Lessmann et al. [5]. On the basis of 124,386 transaction data of the same vehicle model (upper class) of a major car manufacturer, the authors develop a decision support system by means of a support vector regression (SVR). This method extends the classical linear regression such as to permit a non-linear transformation of the independent variable. Each vehicle is described by 176 attributes. The large number of attributes results from the use of dummy variables that represent features such as different optional equipment. Transaction-specific characteristics such as typical features of the customer constitute a crucial point in this method. The authors demonstrate the benefits of using this information in a forecasting model. As such, they recommend the expansion of residual value forecasts within the company to achieve improved predictive power on the basis of exclusive datasets which are not available to external service providers or residual value institutes.

Wu et al. [6] forecast used car prices on the Taiwanese market. Their input parameters comprise the cars brand name, the year of manufacture, the engine type and an equipment index. A new combination of ANNs and ANFIS (adaptive neuro-fuzzy inference system) models are proposed to improve forecasting accuracy. An earlier study carried out by Lian et al. in 2003 describes the problem of the residual value forecast from a time series perspective [7]. Evolutionary Artificial Neural Networks (EANNs) are used in this study to model the residual value of vehicles (all 24 months old) over time (from 1993 to 1997). The authors find cyclical fluctuations, according to which the residual value is at a high level at the beginning of the year and falls to a lower level towards the end of the year (see section 4.1 of our paper for a similar seasonality analysis). Other studies often use mostly macroeconomic indicators in addition to internally available data sources in order to explain the residual value distribution. Prado implements the price of diesel fuel and the industrial production index as explanatory variables in addition to vehicle-specific variables such as age and mileage [8]. At the present time, it is evident that no scientific standard methods are reported in the respective literature. Fan et al. provide a comparison between data mining model approaches such as AutoRegressive Trees (CART), ANNs and linear regression [9]. Their investigations focus on heavy construction machines. According to their analyses, the CART model provides the best results compared to the ANN and the linear regression model.

Besides the purely data-driven forecast techniques, other theoretical model approaches exist to explain price developments and the implications for risk management. This aspect is not a focal point of our work. For a deeper insight, we refer to [10], [11] and [12].

3 System Structure and Input Data

In the following sections, the data and methodology presented in this paper are put in the larger context of a DSS for residual value- and leasing-related issues. In section 3.1, an overview of the complete system structure (see figure 1) is presented, followed by a more in-depth description of the input data used for this study in section 3.2.

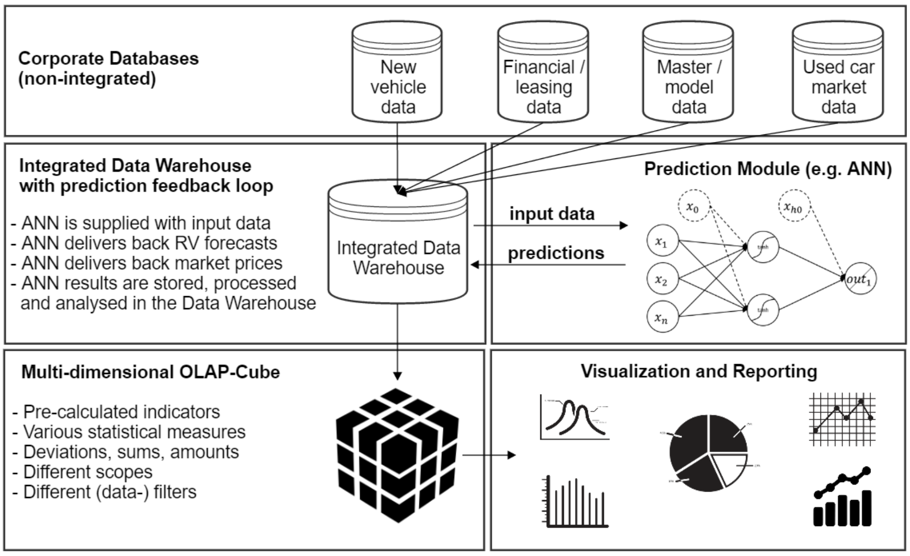


Figure 1. Decision Support System for RV forecasts

3.1 Structure of a DSS for Residual Value Forecasts

In this section, the larger context of the DSS in which the prediction module we focus on is implemented, is explained (see figure 1). The forecasting models need to be trained with historical data. Thus, the first step of the development was to collect and integrate data from different sources within the corporation. The most important of all sources is the used car market database, which stores all used car data collected from the dealerships. It contains every single transacted resale in Germany since 2011, including the Vehicle Identification Number (VIN) as a unique identifier for the vehicle, its mileage, the age of the vehicle, additional model information (e.g. model type, engine, fuel type etc.), resale date and the desired output of the forecasting model, the net resale price. As all data from this source is entered manually in the dealerships and then transferred via a common interface for all dealer management systems in the market, it has to be joined and validated with data from other, internal sources to serve as training data set for the forecasting model (see section 4.2).

One of those sources is the new vehicle database, which collects data from the factory invoicing system directly after production. Besides the corresponding VIN numbers, it contains the original list price of the vehicles, divided into base price, color surcharge and the extra charge for optional equipment, the complete list of optional equipment packages and also an indicator if the vehicle was produced for a promotional offer (with a discount on the original list price).

The information delivered by these systems is decoded using additional data from a Master Data Management system which contains general information about model types, promotional campaigns (such as list price discounts on special edition models), engine power and capacity, gearbox type (manual, automatic, double clutch), number of gears and the model age (number of months since the market launch). Ultimately, there is a connection to the leasing database which provides the initial forecasts and market prices assumed by the captive leasing company. For every single resold vehicle, all of the aforementioned information is gathered, cleansed and stored automatically in a central, integrated data warehouse. This data warehouse, in turn, has a direct connection to the prediction module. The (trained) forecasting model is being fed with input data directly from the data warehouse, predicts market values (actuals) and expected residual values for all vehicles in the database and delivers them back to the data warehouse, where they are stored for further processing and analyses. The next and penultimate layer of the DSS is a multi-dimensional cube for Online Analytical Processing (OLAP). It serves as the basis for visualizations and reports for different purposes, e.g. benchmarking of the forecasting models, market analyses and, most importantly, leasing- and pricing-related decision support. It provides numerous pre-calculated indicators, deviations, sums, amounts, arithmetic means and other statistical measures as well as pre-configured scopes and data filters, e.g. to differentiate between the list price of special edition and standard models or young used cars and cars representing the traditional leasing segment with a leasing term of at least 12 months.

Hence, the OLAP cube can be seen as an extensive toolbox for management reports and visualizations. Used correctly, it can be utilized to find answers for many different used-car- or leasing-related questions, as e.g. individual dealership sales performance rankings, analyses of the used car market price levels after external shocks, monitoring of the car manufacturers' own residual value setting or influences of particular parts of optional equipment on the resale price. These reports need to be configured once on the OLAP level and can then be standardized and visualized with additional software.

The modules previously described, from data collection and integration to reporting and visualization, all play an important role in the operational DSS outlined in this section. Nevertheless, all the results and reports in our particular area of application rely on accurate predictions of actual market values and, even more importantly, most accurate residual value forecasts. Subsequently, the prediction module and its mode of operation represent the core of our DSS and thus, the main focus of the research presented in this paper. The next section 3.2 provides a description of the input data used for the forecasting model.

3.2 Input Data

The data covers a period of four years from 2011 to 2015. Vehicles are included in the database as soon as they have been successfully resold on the used car market after the end of the leasing period. By this means, the residual values realized on the market are included in the database. We distinguish between individual vehicle-specific and model-specific variables. The vehicle-specific variables refer to the leasing contract agreements, which include the mileage, the leasing term (age of the vehicle) and extra charges for special colors and optional equipment. The model-specific variables describe the general characteristics of the vehicle and are the same within each model group. 928 different models are represented in the data. Table 1 shows the available vehicles' features.

Table 1. Input data variables

Variable	*Type*	*Description*
pricecolor	continuous	Extra charge for the color as a percentage of the list price
priceequipment	continuous	Extra charge for optional equipment as a percentage of the list price
mileage	continuous	Vehicle mileage in km/100,000
agevehicle	continuous	Vehicle age (registration date to resale date in days)
pricecolor	continuous	Extra charge for the color as a percentage of the list price
agemodel	continuous	Age of the model since its market launch in months
enginecapacity	continuous	Engine capacity in cubic centimeters
enginepower	continuous	Engine power in horsepower
fourwheel	binary	Four-wheel drive indicator
fuelcode	binary	Gasoline (B), Diesel (D), CNG (EG), LPG (AG)
geartype	binary	Transmission types: automatic (A), double-clutch (D), manual (S)
residualvalue	continuous	Residual value in percent ($\frac{resale\ price}{list\ price}$)

These variables are directly related to the vehicles under investigation. Using the available data, a linear regression analysis is first carried out to explain the residual values. The linear model is specified with all the variables listed in table 1. As expected, the age of the vehicle as well as the mileage have a significant negative impact on the residual value. For a forecasting application, the model age is also a crucial factor. Here we face the problem that the age of a specific vehicle model constantly increases over time. Compared to the age of a specific vehicle where we have a wide range of training examples, the age of the vehicle model is not available beyond the present point in time. The results of this regression analysis indicate that in-sample, the effect of the model age is significantly negative. Using this factor in a forecasting application therefore requires extrapolation techniques. Another interesting finding is that the extra charge for optional equipment has a negative effect on the residual value, which shows that the more optional equipment a vehicle is fitted with, the more value it loses over time in percentage terms.

4 Forecasting Model Specification

4.1 Trend and Seasons in the Used Car Market

In this section we use the residuals of the previously mentioned regression analysis to investigate the phenomenon of seasonality and trend in the used car market, which provides valuable insights to perform a subsequent unbiased forecast. Using the regression approach, all effects of the aforementioned independent variables are eliminated. The following approach is applied: firstly, residuals are sorted with respect to time. Secondly, the residuals for each month during the observation period are averaged. This results in a monthly time series of averaged residuals. Figure 2 shows the resulting time series.

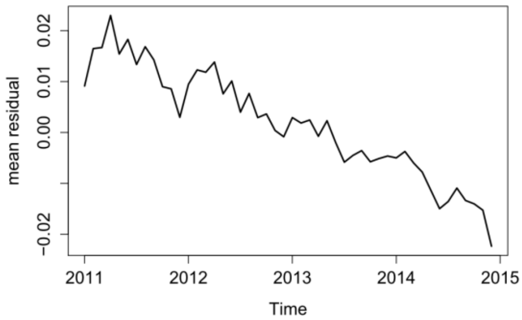


Figure 2. The influence of the time factor on the residual value

The results show a clear downward trend during the four years of observation. After controlling for all contract-specific variables, the results indicate a price decline based on general market conditions (the regression already controls for the constantly increasing model age) for used commercial vehicles. Since we are also interested in seasonal patterns, a decomposition of this time series is performed by the Seasonal and Trend decomposition using Loess (STL) method [13]. A seasonal period of 12 (for each month) is assumed. Figure 3 shows the time series decomposition to determine seasonal and trend effects. The ordinate shows the mean residuals in total (data) and their three different components. It can be observed that the residuals are subject to a trend and a season. The detrended seasonal component shows that towards the end of the year, lower residual values are achieved than during the rest of the year. Accordingly, the residual values tend to be higher during spring, even though these effects are rather small (in the range of 0.8 percentage points of the list price during a year). In order to control for the effects analyzed above, a new regression model is specified according to equation 1, which incorporates a vehicle model independent time factor and also controls for the seasonal components using monthly dummy variables.

$$rv_k = c + \sum_{j=1}^{13}(\beta_j \cdot feature_{j,k}) + \gamma \cdot Time_k + \sum_{m=1}^{11}\left(\delta_m \cdot Month_{m,k}\right) + \varepsilon_k \qquad (1)$$

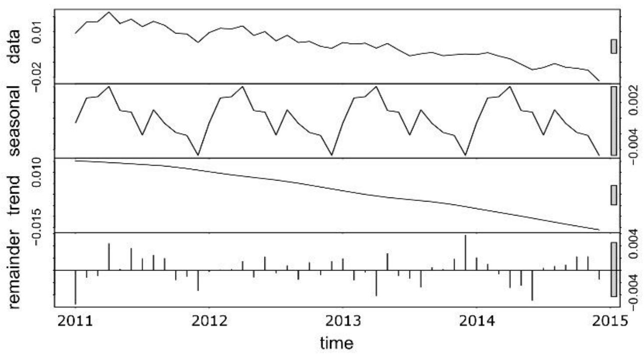


Figure 3. Time series decomposition: seasonal and trend effects

We use the resulting residuals to test for misspecification due to non-linearity with the BDS (Broock, Dechert and Scheinkman) test [14]. The null hypothesis (ε_k is iid) is rejected by the test at the 1% level ($\Delta\varepsilon = \sigma\varepsilon$ and embedding dimension m = 2). Therefore, the linear specification is still rejected by the BDS test. In a forecasting application which can be used in practice, there are tools available to incorporate non-linear relationships in a very effective way. These require careful data handling in comparison to linear models to avoid misleading results. This especially applies to noisy real world data. ANNs are introduced in the following section, which also provides a description of how to preprocess and prepare the data on hand.

4.2 Forecasting with Artificial Neural Networks

ANNs [15] are particularly suitable for the use with non-linear relationships. Their applications range from forecasts in the finance [16], [17] and risk management [18] area to Decision Support Systems [19], [20]. They are robust to very noisy, unstructured, or missing data [21], have powerful pattern recognition capabilities [22] and are therefore well suited to the present real-life problem. Even though many previous studies have shown the potential of this method in forecasting and prediction applications, a proper implementation and validation as well as an accurate data pre-processing is necessary to achieve reliable results [23]. ANNs may be viewed as a method for non-linear function approximation. In this paper, feed forward networks are used. These are composed of several layers of neurons. A first layer (input layer) describes the independent variables which are used to explain the phenomenon. These neurons are connected by weights θ_1 to a further layer (hidden layer), which in turn is connected to the output $h_\theta(X)$ (or a further hidden layer) by the weights θ_2. The hidden layer is thus in turn the input for the following layer. This pattern can be repeated any number of times (any number of hidden layers), each with any number of neurons within the layers. A three-layer feed forward ANN is defined by:

$$h_\theta(X) = \theta_2 \tanh(\theta_1 X) \tag{2}$$

The hidden neurons and, optionally, the output neurons transform their weighted sum of inputs by means of an activation function (usually the hyperbolic tangent).

$$f(X) = \tanh(X) = 1 - \frac{2}{e^{2X}+1} \tag{3}$$

The structure of an ANN is shown in figure 4. To be able to perform a function approximation, the ANNs are trained with training patterns (input x_k, output y_k represent the training pattern k). The randomly initialized parameters (weights) are determined using an iterative process. To avoid the problem of overfitting, a cross-validation is performed. Therefore, the training patterns are split into a set on which the ANNs are actually trained, I_t, and a set of validation patterns, I_v, to estimate the out-of-sample performance. The approximation quality of the ANN is evaluated by calculating the training and validation error functions

$$\varepsilon_t \equiv \frac{1}{2}\sum_{k\in I_t}(h_\theta(x_k) - y_k)^2 \text{ and } \varepsilon_v \equiv \frac{1}{2}\sum_{k\in I_v}(h_\theta(x_k) - y_k)^2. \tag{4}$$

We use an early stopping approach (error on the validation set increases) to obtain ANNs with good generalization capabilities. The best performing ANNs (based on validation error) are then combined in an ANN ensemble, so that the arithmetic mean is used as the actual result of the forecasting model. In this study, the “Fast Approximation performed with Universal Neural Networks” (FAUN) neurosimulator [24] is used for this purpose.

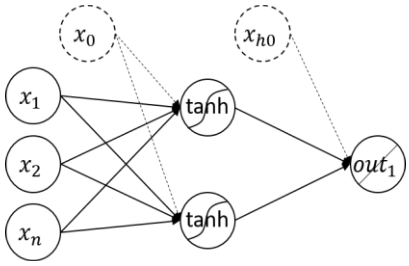


Figure 4. ANN (three-layer feedforward network)

As mentioned above, a usable forecasting application involves constantly increasing, time-dependent parameters: the model age and the time factor in general, which includes all external economic influences and market conditions. Both factors and their influence are not measurable (represented in the data) beyond the present point in time. Including data from the whole time period in the training process (e.g. separating training and testing data randomly from the database) leads to a look-ahead bias in the forecasting result. The training and forecasting periods are presented in figure 5. The forecasting model is trained with data from 2011 and 2012. The out-of-sample performance is documented with data of completed leasing contracts in the following two years (half-year intervals).

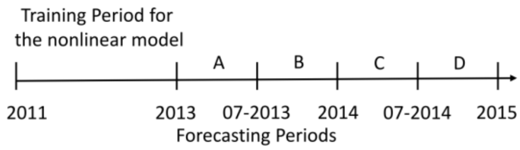


Figure 5. Forecast time line

Since ANNs are sensitive to outliers, we perform a data cleansing before the actual training process is started. For this purpose, 2000 three layer perceptrons including all variables shown in table 1 are trained to a local minimum on the training data. The best ANN (in-sample) is selected and tested for errors. Subsequently, every data set for which $|h_\theta(x_k) - y_k|$ is greater than a threshold value c is removed. As the threshold value c is defined as the 95% quantile of the errors, this means that 5% of the data are removed. This process is repeated a second time. We choose this approach as a result of an extensive cross validation process in which we evaluate the out-of-sample errors of the resulting ANN models with different data cleansing methods (e.g. removing more/less data; performing the procedure only once or three times; manual outlier removal). A manual analysis of the identified outliers reveals that in the case of outliers with higher residual values than expected, additional equipment (like e.g. additional cooling or tool storage equipment) had been installed after specifying the leasing contract, which led to a higher initial value and thus, also to a higher residual value. In the case of outliers with lower residual values than expected, the cars' overall condition (e.g. accidents) had not been properly reported in the data filtered by this approach. As we aim for accurate residual value predictions for accident-free vehicles without additional installations, these outliers had to be excluded from our analyses. For the actual training of the forecasting model, three-layer perceptrons with 4, 6 or 8 hidden neurons were tested. For each of the different topologies, 2000 networks were randomly initialized and trained to a local minimum. On the basis of validation data which is not represented in the training data set, the best 30 networks were selected and the arithmetic means of these approximations were used as a result. The ANN models with 6 hidden neurons yielded the smallest error regarding the validation data. In the next section, the results of the analyses are presented.

5 Results

This section presents the results of the linear and non-linear function approximations. ANNs generally provide good results for interpolation tasks. This means that the function approximation works well if the presented values lie within the input/output ranges of the training data. An extrapolation far beyond these ranges is difficult to realize. The forecasting application with ANNs for the third and fourth year (hence not represented in the training data regarding the time variables) uses the characteristics of new vehicles as inputs but sets the model age and time factor to the maximum value represented in the available data for these periods. The result is a so-called current market value, which reflects actual prices on the used car market. Table 2 (a) shows the resulting bias (mean error defined in equation 5) when using these forecasting models for predictions in the following two years.

$$mean\ error = \frac{1}{n}\sum_{i=1}^{n}(real\ value_i - forecast_i) \tag{5}$$

Table 2. Adjusted and non-adjusted forecasts of the ANN

(a) Non-adjusted forecast

Period	*Mean error*	*P-value*
A	0.00836	0.0524
B	-0.01028	0.0074**
C	-0.01551	<0.001***
D	-0.04640	<0.001***

(b) Adjusted forecast

Period	*Mean error*	*P-value*
A	0.00836	0.0524
B	0.00220	0.5637
C	0.00533	0.1817
D	-0.01787	<0.001***

***, **, * indicate statistical significance at the 0.1%, 1% or 5% level, respectively.

A t-test with the null hypothesis that the mean error is zero (unbiased) already rejects the hypothesis in the second forecasting period at the one percent significance level.

The actual residual value (forecast), in contrast to the market value, is now adjusted by a linear factor resulting from a time series regression of the first two years of data. The adjustment of the ANN forecast is performed according to equation 6.

$$rv_{t+n} = Market\ Value_t + \left(\beta_{agemodel} + \beta_{Time}\right) \cdot n \tag{6}$$

Table 2 (b) shows the results after the linear adjustment. The adjusted results show that the bias is now also insignificant for the forecasting periods B and C. For period D, the forecasting model is still biased. An investigation of possible non-linear dependencies in the time dimension is a topic for future research. This implies that a database of more than one complete business cycle must be investigated. The linear adjustment currently undertaken is a correct assumption for at least one and a half year forecasts, as indicated by the time series decomposition shown in figure 3 and the results listed in table 2 (b). The proposed linear model specified by equation 1 is now tested against the ANN forecast with a linear adjustment for the model age and time factor. To evaluate the linear model in a forecasting benchmark, a regularization term is incorporated for the training process. We choose a ridge regression approach (L2 norm) to prevent potential overfitting of the linear model. The lambda parameter is selected by cross validation. Due to the small number of parameters compared to the large number of training patterns in the years 2011 and 2012, the lambda parameter with the best performance is close to zero (0.04), so the linear model barely suffers from a high variance problem. Table 3 shows the results, the root-mean square error (RMSE) and the mean-absolute error (MAE), of the two forecasting methods for each of the four time periods. As evident from the increasing trend in the error measures, the forecast becomes more inaccurate the longer the forecasting period is. The non-linear relationships between the vehicle characteristics are represented significantly better by an adjusted ANN method than by a pure linear model. The Diebold-Mariano test [25] rejects the hypothesis that both methods have the same accuracy at any confidence level.

Table 3. Benchmark results

Period	*Measure*	*Adjusted ANN*	*Linear*	*DM-Test*
A	RMSE	0,07834	0.08554	<0.001***
	MAE	0.05638	0.06287	
B	RMSE	0.08026	0.08824	<0.001***
	MAE	0.05821	0.06515	
C	RMSE	0.07995	0.08965	<0.001***
	MAE	0.05891	0.06788	
D	RMSE	0.08218	0.09113	<0.001***
	MAE	0.06051	0.06857	

***, **, * indicate statistical significance at the 0.1%, 1% or 5% level, respectively.

Besides the significant performance improvements, these results are indeed economically relevant in our practical application compared to former, simplistic predictions based on data from external service providers for residual value estimations. In the next section, we discuss the results and provide an outlook for further research.

6 Discussion

The challenge of forecasting residual values for commercial vehicles was examined in this paper. An accurate forecast as a substantial part of a larger DSS is crucial for managing the exposure risk of car manufacturers and dealers in the leasing business. A systematic bias in both directions (too high or too low predictions) leads to negative consequences, either for competitiveness or the resale margin. The aim was to establish a purely data-driven approach for forecasting residual values based on information about past transactions. The results indicate that ANNs are well-suited for such noisy and unstructured data. Nevertheless, thorough data preparation and outlier detection is important in order to achieve good results. We show how ANNs can be used to preprocess and filter the database and how a forecasting model can be designed based on an ANN approach. ANNs are often criticized owing to their black-box nature. By means of a transparent description of the topology selection and data cleansing process, it is possible to mitigate the problem and make the results reproducible. ANNs are often only used at present as an alternative tool to benchmark different methods. However, an optimization within the whole ANN framework can help to improve the forecasting capability of this method.

Accordingly, the influence of the model age on the residual value over time and a general time factor have been investigated in more detail. This is necessary due to the fact that the model age (the period of time a specific vehicle model has been on the market since its market launch) is a factor which is not measurable beyond the present point in time, as it is a continuously increasing value. This also applies to the time factor, which takes account of all external market conditions. Although both factors remain uncertain, they play a crucial role in an actual, usable forecasting application. All other explanatory variables such as the age of a specific vehicle or its mileage are represented in the data for a sufficiently wide range of values. In the present study, the time-dependent factors were investigated in detail in order to linearly adjust the non-linear model to mitigate this problem. A clear downward trend in the residual values was identified with increasing model age and time. A reason for this may be the declining attractiveness of a particular vehicle model, which means that customers tend to buy a newer generation of products. As a result, older vehicle models must be offered at a lower price.

Moreover, the seasonal component of the used car market shows systematically lower residual values in December and higher residual values during the first quarter of a year. This empirical observation may be explained, e.g. by dealers trying to meet their sales targets towards the end of the year. Although these observations might be a special phenomenon for the type of vehicles examined in our study (commercial vehicles), it is a topic for further research investigating these patterns for a broader range of vehicle classes in order to formulate a more general statement. The entire analysis carried out in this paper is based on data between 2011 and 2015. From an economic point of view and in relation to the automotive market, this period does not include any major crises or boom phases. As the DSS and all necessary interfaces presented in section 3 are already successfully implemented, market data, as well as the corresponding internal data from corporate systems and databases, is being automatically recorded and stored in the central data warehouse for further processing. It is of interest to investigate how the results and the forecasting model quality change by analyzing data covering a longer period of time, including information about the behavior of residual values during crisis and/or boom phases in the database.

As previously mentioned, it is possible that external factors such as consumer demand, competitors' decisions, fuel prices or macroeconomic developments, e.g. measured by an industrial production index, also have a significant impact on the achievable resale price. Such factors were not included in this study for two reasons. Firstly, a complete theory which explains all factors that influence the used car market does not exist. This could take the form of a "fishing license" (alluding to the factor models of asset pricing). Up to now, however, not enough data is available to test the out-of-sample influences with a sufficiently long time horizon. Secondly, the problem of a look-ahead bias is the major concern in this application. The factors which drive prices on the used car market must be measurable at the time of the contract conclusion. This is a necessary precondition for performing a usable forecast in practice. In order to avoid biased predictions, however, economic indicators must also be forecasted to provide additional forecast model input data. To achieve a broader understanding of residual value data in general, an extensive study about different kinds of internal and external factors must be conducted in addition to benchmarks of proposed forecasting methods in the literature. Since each of the aforementioned, related studies have their own protected data with different characteristics, a benchmark of their methods is hardly interpretable if it is performed solely on one specific data set, because one method may not suit all kinds of data. Addressing this issue is an important topic for further research.

The above paragraph of the discussion is directly focused on the forecasting methodology implemented in our system. In the greater context of an entire DSS, the particular use case presented in this paper serves as a general example that by using methods and tools from the field of business intelligence, predictive analytics and data science, corporations can obtain valuable, business-critical information from data that is often already internally available. It also shows that this information, properly prepared and visualized, can support or even induce management decisions and help to monitor the consequences. In the particular case of the DSS presented in section 3, many use cases beyond residual value forecasts have already been realized, from the assessment of individual dealership performance to the measurement of the impact of internal decisions or external effects. Another practical example from our project is the simulation of different market scenarios, e.g. the assessment of the car manufacturers' risk exposure if residual values suddenly drop due to an unexpected external shock such as financial crises, new legislations or competitors' decisions. The results of such "stress tests" support the financial/controlling department in the establishment of appropriate provisions to cover these risks.

One of the major disadvantages of such systems is that they are usually quite complex. Understanding the whole process of how results are determined requires a deep technical and mathematical understanding and, most importantly, the necessary time. At this point, complexity may become a problem, as time usually is a scarce factor in management. Decision makers need to be able to trust the results, especially if important decisions are to be based on them. In an operative, real-life DSS, it is our experience that technology acceptance can be strengthened e.g. by a clearly arranged data quality assurance dashboard and, in general, reports whose visual appearance is not completely new to the receivers' eye. Subsequently, professional communication of project results, analyses and ultimately, between managers, analysts and developers is inevitable and plays an important role in earning the necessary confidence.

7 Conclusion

In the course of this paper, we investigated means to reduce the residual value risk car manufacturers and leasing companies face in their daily business. For this purpose, a DSS with forecasting capabilities was developed, implemented and tested. As accurate residual value forecasts are a critical success factor in this business, the focus has been put on suitable forecasting methodologies. ANNs have proven to be a reliable method for this purpose and thus, were implemented in the forecasting module of the DSS presented in this paper. It is necessary to mention that this system is in operative use and under constant improvement, as data is being automatically collected, cleansed, stored and used to improve the accuracy of future residual value forecasts. Furthermore, the flexibility of the integrated data warehouse and OLAP cube allow the investigation of a wide range of further, business-critical issues. From the performance measurement of dealerships to the impact of pricing strategies, promotions and discounts on residual values as well as interdependencies of the new and used car market, many future research opportunities are conceivable.

From a more general perspective, it can be concluded that corporations usually have access to large amounts of exclusive, unstructured, unintegrated data, which, properly processed, prepared and interpreted, have the potential to make a difference in daily business or even in important strategic decisions. This potential should not remain unexploited. Nevertheless, the information obtained still relies on humans with expert knowledge to interpret and communicate the results correctly. In a practical environment, close coordination between these experts and decision makers is crucial for the success of such projects. Since easy to understand reports can reduce complexity, build up trust and thus, facilitate communication, further research should also concentrate on proper visualization of data and analyses.

References

1. Leaseurope, 2014 leasing key facts & figures. http://www.leaseurope.org (Accessed: 19.07.2016)
2. DAT Group, DAT Report 2016. https://www.dat.de/report (Accessed: 16.07.2016)
3. Prado, S., Ananth, R.: Breaking through risk management, a derivative for the leasing industry. Journal of Financial Transformation, 34:211–218 (2012)
4. Nelson, M., Hill, T., Remus, W., O'Connor, M.: Time series forecasting using neural networks: should the data be deseasonalized first? J. of Forecasting, 18(5):359-367 (1999)

5. Lessmann, S., Listiani, M., Voß, S.: Decision Support in Car Leasing: A Forecasting Model for Residual Value Estimation. In: M. Lacity, F. Niederman, S. March (eds.) Proc. of the Intern. Conf. on Information Systems, Saint Louis, MO, USA (2010)
6. Wu, J. D., Hsu, C. C., Chen, H. C.: An expert system of price forecasting for used cars using adaptive neuro-fuzzy inference. Expert Systems with Applications, 36(4):7809–7817 (2009)
7. Lian, C., Zhao, D., Cheng, J.: A fuzzy logic based evolutionary neural network for automotive residual value forecast. In: Proceedings of the International Conference on Information Technology: Research and Education, pp. 545-548 (2003)
8. Prado, S.: The European used-car market at a glance: Hedonic resale price valuation in automotive leasing industry. Economics Bulletin, 29(3):2086- 2099 (2009)
9. Fan, H., AbouRizk, S., Kim, H., Zaïane, O.: Assessing residual value of heavy construction equipment using predictive data mining model. Journal of Computing in Civil Engineering, 22(3):181–191 (2008)
10. Rode, D. C., Fischbeck, P. S., Dean, S. R.: Residual risk and the valuation of leases under uncertainty and limited information. Journal of Structured Finance, 7(4):37–49 (2002)
11. Storchmann, K.: On the depreciation of automobiles: An international comparison. Transportation, 31(4):371–408 (2004)
12. Smith, L. D., Jin, B.: Modeling exposure to losses on automobile leases. Review of Quantitative Finance and Accounting, 29(3):241–266 (2007)
13. Cleveland, R. B., Cleveland, W. S., McRae, J. E., Terpenning, I.: STL: A seasonal-trend decomposition procedure based on loess. Journal of Official Statistics, 6(1):3-73 (1990)
14. Broock, W. A., Scheinkman, J. A., Dechert, W. D., LeBaron, B.: A test for independence based on the correlation dimension. Econometric Reviews, 15(3):197-235 (1996)
15. Bishop, C. M.: Neural networks for pattern recognition. Oxford University Press (1995)
16. Sermpinis, G., Stasinakis, C., Dunis, C.: Stochastic and genetic neural network combinations in trading and hybrid time-varying leverage effects. Journal of International Financial Markets, Institutions and Money, 30:21–54 (2014)
17. Zimmermann, H., Neuneier, R., Grothmann, R.: Multi-agent modeling of multiple FX-markets by neural networks. IEEE Transactions on Neural Networks, 12(4):735–743 (2001)
18. von Spreckelsen, C., von Mettenheim, H.-J., Breitner, M. H.: Real-time pricing and hedging of options on currency futures with artificial neural networks. Journal of Forecasting, 33(6):419–432 (2014)
19. Eilers, D., Dunis, C. L., von Mettenheim, H.-J., Breitner, M. H.: Intelligent trading of seasonal effects: A decision support algorithm based on reinforcement learning. Decision Support Systems, 64:100-108 (2014)
20. Kuo, R. J., Chen, C. H., Hwang, Y. C.: An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network. Fuzzy Sets and Systems, 118(1):21–45 (2001)
21. Schocken, S., Ariav, G.: Neural networks for decision support: Problems and opportunities. Decision Support Systems, 11(5):393–414 (1994)
22. Zhang, G., Patuwo, B.E., Hu, M.Y.: Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, 14:35–62 (1998)
23. Adya, M., Collopy, F.: How effective are neural networks at forecasting and prediction? A review and evaluation. Journal of Forecasting, 17(5-6):481-495 (1998)
24. Mettenheim, H.-J., Breitner, M.: Robust decision support systems with matrix forecasts and shared layer perceptrons for finance and other applications. Proceedings of the International Conference on Information Systems, Saint Louis, MO, USA (2010)
25. Diebold, F. X., Mariano, R. S.: Comparing predictive accuracy. Journal of Business & Economic Statistics, 13(3):253–263 (1995)

Disruption of Individual Mobility Ahead? A Longitudinal Study of Risk and Benefit Perceptions of Self-Driving Cars on Twitter

Christopher Kohl[1], Dalia Mostafa[1], Markus Böhm[1], Helmut Krcmar[1]

[1] Technical University of Munich, Department of Informatics, Munich, Germany
{christopher.kohl,dalia.mostafa,markus.boehm,krcmar}@in.tum.de

Abstract. In this paper, we address the question if there is a disruption of individual mobility by self-driving cars ahead. In order to answer this question, we take the user perspective and conduct a longitudinal study of social media data about self-driving cars from Twitter. The study analyzes 601,778 tweets from March 2015 to July 2016. We use supervised machine learning classification to extract relevant information from this huge amount of unstructured text. Based on the classification, we analyze how risk and benefit perceptions of self-driving cars develop over time, and how they are influenced by certain events. Based on the perceived risks and benefits, we draw conclusions for the acceptance of self-driving cars. Our study shows that a disruptive innovation of self-driving cars is not likely as risk and benefit perception issues indicate a lack of acceptance. We provide suggestions for improving the acceptance of self-driving cars.

Keywords: Machine learning, Risk Perception, Self-Driving Cars, Technology Acceptance, Text Classification

1 Introduction

In this paper, we address the question if there is a disruption of individual mobility by self-driving cars ahead of us. The impressive recent technical developments, for example of the Google Car and the Tesla Autopilot, draw a performance trajectory characteristic for disruptive innovations [1]. They already demonstrate the technical feasibility of self-driving cars. However, other previously new technologies in the individual mobility sector such as electric cars [2] or ridesharing [3] have been available since decades but still have a low market share. So will there be a disruption of individual mobility from human-driven cars to driverless cars as it occurred from horse-drawn carriages to horseless carriages as some articles predict [4]?

The evolution of transportation has faced numerous trials as it grew over time. We have gone through many diverse phases, including walking, biking, horses, coaches, trains, and cars. It is safe to assume that this steady chain of development of faster vehicles with improved features continues. Over the past decade, a countless amount of research has been invested into self-driving cars [5]. Companies such as Google,

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Kohl, C.; Mostafa, D.; Böhm, M.; Krcmar, H. (2017): Disruption of Individual Mobility Ahead? A Longitudinal Study of Risk and Benefit Perceptions of Self-Driving Cars on Twitter, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1220-1234

Tesla, and BMW are investing in the development of self-driving cars. Especially because of these high investments, we must remember a significant key factor for the success of emerging technologies: technology acceptance [6].

In recent times, self-driving cars have become a controversial topic (e.g., because of ethical concerns [7]). Despite the efforts of researchers in pushing the technical boundaries of science and technology, there are key factors that need to be considered. One of the most meaningful factor is people's concerns regarding this emerging technology [8]. People's perceived risks and benefits towards self-driving cars will be central determinants of their public acceptance [9]. Public acceptance is what will eventually determine, when and how self-driving cars will actually be put to use, making it a crucial factor to take into consideration. As Michael Toscano, CEO of the Association for Unmanned Vehicle Systems International once said "The technology maturation is there, but the public acceptance is not there" [10].

Opinions regarding self-driving cars such as risk and benefit perceptions are affected, and perhaps even shaped, by the news [11]. If we succeed in explaining the logic behind people's various opinions concerning self-driving cars, we will be one step closer towards tackling the issue of technology acceptance. Therefore, we use supervised machine learning classification to extract this information from a set of 601,778 tweets obtained from the microblogging service Twitter.

Twitter has often proven to be a valuable source of data for prediction and monitoring of diverse phenomena ranging from disease outbreaks [12] to political elections [13]. Users of Twitter face a limit of 140 characters per message, referred to as "tweet", to include all relevant information. Despite their brevity, tweets contain valuable information encoded in natural language [14]. It is an ongoing challenge to extract this information from the vast amount of noise present on Twitter. We build on previous findings from sentiment analysis [14] and machine learning classification to extract information from a rich dataset of tweets.

The remainder of this paper is structured as follows. First we give an overview about technology acceptance literature and self-driving cars in general in section 2. Second, we describe the data extraction from Twitter, preprocessing the data, and model generation including its evaluation in section 3. Third, we describe the results in section 4. Fourth, we discuss our results in section 5. In section 6, we conclude with a summary of the results, limitations, possibilities for further research, and contributions to research and practice.

2 Theoretical Background

In this section, we give an overview about current literature disclosing the significance of acceptance towards self-driving cars from an Information Systems (IS) and public acceptance perspective. We give an introduction to self-driving cars and present the current scientific knowledge and surveys relevant to the acceptance of self-driving cars. We conclude this section by summarizing the theoretical background, thereby motivating the research from a theoretical perspective.

2.1 Technology Acceptance

Technology acceptance is one of the main research streams of IS research and the technology acceptance model (TAM) being a crucial source of many research endeavors [15]. The aim of TAM is to explain and predict if and why information systems will be used by individuals [6]. TAM predicts user acceptance by using three basic constructs: Perceived usefulness, perceived ease of use, and behavioral intention to use the system under consideration.

Several models were derived from the TAM with the Unified Theory of Acceptance and Use of Technology (UTAUT) being one of the most established ones that integrates eight models of technology adoption including TAM [16]. It includes the constructs of TAM and adds social influence (i.e., the degree to which influential people think the user should use the particular system) and facilitating conditions (i.e., the perceived level of organizational and technical support for the system, which is also considered a direct predictor of technology use). Individual factors such as age and gender moderate the relationships between these constructs and technology acceptance and use. Several researchers have extended the UTAUT model [17].

Many extensions of TAM and UTAUT have recognized the importance of risk perception for user acceptance. For example, Martins et al. [18] study Internet banking adoption and conclude that risk perception is an important factor. Lancelot Miltigen et al. [19] study end-user acceptance of biometrics and find that the greater the perceived risks, the lesser people will accept this technology. Despite several promising approaches, risk perception has not been included in one of the central IS acceptance models [17].

Public acceptance research recognizes that many technologies have been rejected by people because of societal controversies, causing negative consequences for the commercialization of technologies [8]. Considering the vast investments in research and development of self-driving cars and the potential benefits of this technology for society, rejection of this technology could have severe consequences. In particular, unpredicted events and accidents that recently occurred with self-driving cars such as the first human casualty [20] could lead to fear and reluctance to adopt.

A very influential model of technology acceptance in the public acceptance field specifically focuses on the relationship between perceptions of risks and benefits, trust, and technology acceptance [9]. The study found that perceptions of risks and benefits directly influence technology acceptance.

2.2 Self-Driving Cars

The National Highway Traffic Safety Administration (NHTSA) [21] defines five degrees of car autonomy which have different extents of connection between cars and the Advanced Driver Assistance Systems (ADAS) and the level of control the car carries. These systems can have full control of the car or can just be an assistance system for the driver. The levels vary from non-autonomous at all to fully-autonomous and are defined as follows [21]:

- Level 0: (Non-autonomous): The driver is in complete control of the vehicle.

- Level 1: (Function Specific Automation): Automation involves only specific control functions. (i.e. pre-charged breaks, electronic stability control)
- Level 2: (Combined Function Automation): Automation of two primary control functions in unison to relieve driver of control of these functions.
- Level 3: (Limited Self-Driving Automation): The driver has the choice to give up control of all safety-critical functions under certain conditions, yet the driver is expected to be available for occasional control.
- Level 4: (Full Self-Driving Automation): The vehicle has full control of all safety-critical driving functions under all conditions. The driver's availability is completely unnecessary.

The current automation level of self-driving cars is level 2. The drivers are still required to monitor the car and need to be ready to take over control at any time. There could be severe consequences if a driver fails to comply (e.g., [20]). However, many drivers are misusing the system, for example by even leaving the driver's seat entirely while driving on a public road using the Autopilot feature of a Tesla Model S [22]. Considering how difficult it is for the driver to get back in the loop and react properly to certain traffic situations [23], such reports are even more troubling and show that also exaggerated benefit perceptions could have negative implications for technology acceptance.

Recent surveys have indicated that 56% of people have positive opinions towards self-driving cars, while 13.8% carry negative concerns, and 29.4% are neutral towards the topic [24]. Supporters argue that since 93% of car accidents are due to driver error [25], the use of self-driving cars would reduce car accidents by that exact amount [5]. However, opponents of this view state that these vehicles would introduce new risks that do not exist now, such as system failures or offsetting behaviors. Schoettle and Sivak's analysis [24] concluded that self-driving cars may be no safer than an average driver and that they may result in the increase of total crashes if self- and human-driven vehicles are used simultaneously.

Many recent surveys have shown that people are generally accepting self-driving cars (e.g., [26]) even if only little is known about the technology. If self-driving cars become available people may just begin to recognize potential issues as it was the case with active cruise control where people began to recognize the loss of control at the first time deployment [27].

2.3 Summary

Risk and benefit perceptions are likely to play a central role for the acceptance of self-driving cars. Even before public availability, risk and benefit perceptions should be closely monitored to identify the issues of people with the technology. Issues can be accurate risk perceptions that need to be addressed or benefits that can be exploited in an early stage of development. Extensions of the TAM, UTAUT, and models from other fields of research have shown that risk perceptions are direct antecedents of technology acceptance.

Another kind of issues are distorted perceptions of both benefits and risks [28], which we already see with the first available self-driving car technologies. An overestimation of benefits might lead to misuse of self-driving cars, disappointment of initial users, and can have fatal consequences. Underestimation or not even recognizing benefits on the developer side could lead to self-driving cars that do not exploit their full potential. An overestimation of risks by the public could lead to resistance against self-driving cars before they even become publicly available [29].

Taking this into account, we identify the need to study risk and benefit perception of self-driving cars. Instead of distributing questionnaires, we use a novel approach to identify risks and benefits by analyzing the vast amount of existing data about self-driving cars on social media. We use supervised machine learning classification to classify tweets, which allows us to analyze them qualitatively and quantitatively. Classification of documents written in natural language is a common approach from opinion mining [30]. Thereby, we avoid certain issues with questionnaires and studying technology acceptance, for example common method variance [31].

3 Method

In this section we describe our approach from data extraction to model application. We follow the process suggested by [32]. First, we obtain tweets using the Twitter Search API. Second, we preprocess the tweets to improve data quality, reduce dimensionality, and avoid misclassification. Third, we evaluate the machine learning classification algorithm.

3.1 Data Extraction

The dataset consists of tweets concerning self-driving cars that were obtained using the Twitter Search API [33]. Furthermore, we developed a Java application as the Twitter Search API only allows to retrieve tweets not older than one week [34]. In order to conduct a meaningful longitudinal analysis, it was essential to allow for longer date intervals by fetching the tweets daily and storing them in a database. A MongoDB NoSQL database was used to store the complete tweets as they were returned by the Twitter API including their date of creation, the username of the tweet creator, the message that was tweeted, and a unique identifier of the tweet. We started the data collection for this analysis on March 03, 2015 with the last tweets being posted on July 15, 2016. We used the following set of search queries (SQ) in our Twitter API requests:

- SQ1: self driving OR driverless OR autonomous OR automated
- SQ2: tesla OR google OR apple OR icar OR ford OR opel OR gm OR general motors
- SQ3: volkswagen OR vw OR daimler OR mercedes OR benz OR bmw OR audi OR porsche

The search queries have been fixed before the data collection and consist of a combination of topic-related keywords (SQ1), names of U.S.-based companies working on self-driving cars (SQ2), and German car manufacturers (SQ2 and SQ3). Especially

SQ2 and SQ3 resulted in many tweets that were not concerned with self-driving cars. However, at the beginning of our research in March 2015, we wanted to make sure that the search queries still find the relevant tweets without having to change the search queries. In total, we collected 1,859,619 tweets. For the data analysis, the tweets were filtered using a regular expression[1], which ensures that only tweets containing one of the following terms are included in the data analysis: driverless, self-driving, autonomous driving, automated driving, autonomous car, and automated car. In addition to traditional filtering using strings, the regular expression also allows slight variations of the terms, such as "driver less" or "driver-less". This selection method reduced the number of tweets to 601,778.

For training the machine learning classifier we used a dataset of 7,482 tweets, which were manually classified by one person using the three labels "Risk", "Benefit", and "Neutral". "Risk tweets" describe perceived risks of self-driving cars while "Benefit tweets" describe benefit perceptions of self-driving cars. "Neutral tweets" do not contain risk nor benefit perceptions, for example: "Google starts testing driverless car in Austin […]" or "New self-driving Google car heads to streets […]." The distribution of the tweets is shown in Table 1.

Table 1. Descriptive statistics of the training dataset

		Class	
	Risk	*Benefit*	*Neutral*
N	751	701	6,030
%	10.0	9.37	80.6

The tweets were created in the time range from beginning of January 2010 to June 2014 and collected by crawling the "top tweets" about self-driving cars from the Twitter website prior to this study. These are "popular Tweets that many other Twitter users have engaged with and thought were useful" [35]. Both, the training dataset and the collected tweets were created by potential consumers and from users with commercial interests, for example, self-driving car manufacturers or news providers. For this analysis, we will not differentiate between the authors of the tweets. With the "top tweets", we could get an overview of the discussion about this topic on Twitter, which helped to design this study. However, we refrain from analyzing these tweets since they only represent a small fraction of the actual tweets published from January 2010 to June 2014 and are probably highly biased through the proprietary selection algorithms of Twitter. We only use them as "training data" for machine learning classification.

3.2 Data Preprocessing

We performed changes to the content of the tweets to reduce dimensionality and avoid misclassification, which is a common step in text classification [32]. We use the text

[1] We used the following regular expression: (driver.?less | self.?driving | autonomous.?driving | automated.?driving | autonomous.?car | automated.?car)

mining package “tm” for preprocessing, which provides a text mining framework for the statics software R [36]. The preprocessing steps are described in more detail in [36].

First, we transformed all characters in the text of the tweets to lower case. Like most of the preprocessing steps, this decreases readability for humans. However, machine learning classifiers for text classification mainly rely on statistical features of the provided textual data and, thus, profit from such transformations. Second, we removed punctuation, numbers, and hyperlinks. Since we will not perform a grammatical analysis, punctuation is not required to determine the classification of the tweets. Third, we remove English stopwords as provided by the tm package. Additionally, we removed the Twitter-specific stopwords “via” and “rt”. Fourth, we use stemming to further reduce dimensionality of the tweets. Stemming reduces words with the same stem to the same word by stripping derivational and inflectional suffixes, for example: “driving” is stemmed to “drive”.

Having performed the described transformations, the text of the tweets now should mainly contain words that are useful for the machine learning classification. In the last step, we transform the textual representation of the tweets into a document term matrix. Only words containing at least two characters and occur at least ten times in the tweets are included as terms. The terms are weighted by the term frequency (i.e., the number of occurrences of a certain term). Terms are, in our analysis, single words (i.e., unigrams) We apply all of the described preprocessing steps to both the training data and the tweets we want to classify.

3.3 Model Generation and Evaluation

The basic idea of supervised machine learning text classification is to automatically assign classes to documents using a much smaller set of training data. The training data usually contains manually classified documents from which the machine learning algorithms create a model that determines how to classify new documents. There are many different machine learning algorithms available for this task such as Naïve Bayes, maximum entropy classification, or Support Vector Machines (SVM) [37].

We decided to use the SVM algorithm for text classification, which has been shown to be highly effective for this task [37, 38]. It does not require extensive parameter tuning and is able to cope well with large feature vectors as it is usually the case with text classification [38]. The basic idea of SVM is to find a hyperplane that separates the documents (i.e., tweets) according to their classification with a margin that is as large as possible, which is basically an optimization problem [37]. We use the LIBSVM implementation of SVM that allows classification, regression, and other learning tasks [39]. For our analysis, we use C-support vector classification for classification.

For this analysis, we set the regularization parameter C to of the SVM to one and select a linear kernel function since text classification problems are often linearly separable [38]. We compute several metrics to evaluate the SVM. First, we conduct a 10-fold cross validation to determine the accuracy of the classifier. Accuracy is defined as the overall number of correct classifications divided by the number of instances in the dataset and a k-fold cross-validation randomly splits the training data into k mutually exclusive, approximately equal sized subsets (i.e., folds) [40]. The algorithm uses one of the *k* folds to evaluate the classifier by computing the accuracy and the other *k – 1* folds to train it. The cross-validation showed an average accuracy of 87.7%, which is a very good value considering similar studies (e.g., [41]) and is much better than classification based on hand-picked keywords [30].

For the second evaluation, we split the training data using a random selection of 80% (N = 5,957) of the tweets for training the SVM and 20% (N = 1,525) for evaluating the classification performance. We then compute several metrics based on the confusion matrix shown in Table 2.

Table 2. Confusion matrix of the SVM algorithm

		True class		
		Risk	*Benefit*	*Neutral*
Predicted class	Risk	80	9	20
	Benefit	4	76	22
	Neutral	61	62	1191

The accuracy with the fixed training set is 88.33%. We computed the "no-information rate", the largest proportion of the observed classes, since there is a large imbalance between the classes [42]. The no-information rate has a value of 80.85%. Additional metrics were computed according to [42] and are listed in Table 3.

Table 3. Metrics by class

Metric	*Risk*	*Benefit*	*Neutral*	*Average*
Sensitivity	0.5517	0.5170	0.9659	0.6782
Specificity	0.9790	0.9811	0.5788	0.8463
Pos. Pred. Value	0.7339	0.7451	0.9064	0.7951
Neg. Pred. Value	0.9541	0.9501	0.8009	0.9017
Prevalence	0.0951	0.0964	0.8085	0.3333
Detection Rate	0.0525	0.0498	0.7810	0.2944
Detection Prevalence	0.0715	0.0669	0.8616	0.3333
Balanced Accuracy	0.7654	0.7491	0.7724	0.7623

While accuracy showed very good values, we could identify issues of the SVM classifier resulting from the imbalanced training set. For example, the difference in sensitivity between Risk and Benefit tweets suggests, that the SVM recognizes benefit-related tweets better than risk-related tweets.

4 Results

With an overall total of 601,778 tweets, we obtained 459.751 (76.4%) neutral tweets, 63,599 (10.6%) stated benefits (*BT*), and 78,428 (13.0%) stated risks about self-driving cars (*RT*). The risk ratio (RR) and benefit ratio (BR) were calculated as follows:

$$RR = \frac{RT}{RT+BT} = \frac{78{,}428}{78{,}428+63{,}599} = 1 - BR = 0.5522 \quad (1)$$

$$BR = \frac{BT}{RT+BT} = \frac{63{,}599}{78{,}428+63{,}599} = 1 - RR = 0.4478 \quad (2)$$

In 2015, we collected 490,284 tweets of which 376,923 (76.9%) of the tweets were neutral, 50,098 (10.2%) stated benefits, and 63,263 (12.9%) stated risks about self-driving cars. The *RR* in 2015 is 0.5581 and *BR* is 0.4419. The ratio of neutral tweets did not change much over the years: Of 111,494 tweets in 2016, 82,828 (74.3%) of the tweets were neutral, 13,501 (12.1%) stated benefits, and 15,165 (13.6%) stated risks about self-driving cars. *RR* in 2016 is 0.5290 and BR is 0.4710. The results are summarized in Table 4.

Table 4. Number of tweets per year by class

Year	*Total*	*Neutral*	*Benefit*	*Risk*	*RR*	*BR*
2015	490,284	376,923 76.9%	50,098 10.2%	63,263 12.9%	0.5581	0.4419
2016	111,494	82,828 74.3%	13,501 12.1%	15,165 13.6%	0.5290	0.4710
Overall	601,778	459,751 76.4%	63,599 10.6%	78,428 13.0%	0.5522	0.4478

The ratio of neutral tweets, *RR* and *BR* did not change much over the years. This could indicate that the SVM classifier and the underlying training data is well-suited for classifying tweets about the risk and benefit perceptions of self-driving cars. It might also show that *RR* and *BR* is a good measure to analyze risk and benefit perception in further research. Closer inspection of *RR* and *BR* showed that it did change between the months (Figure 1) and might be an important indicator for issues in risk and benefit perception. However, as the SVM classifier detects benefit-related tweets better than risk-related tweets, the *RR* (*BR*) metric is suspected to be lower (higher) than the reported one.

We identified a spike in *BR* in August 2015 in Figure 1. By inspecting the tweets from August 2015, we found that many tweets mentioned the announcement of autonomous crash trucks that help to improve safety at road construction sites [43]. Drivers of crash trucks are usually in a very dangerous situation. Removing the driver could save many lives and was obviously very well received by the public.

Plotting the tweets over time, we could observe several changes in the number of risk and benefit tweets. For example, the graph of risk tweets (Figure 2) shows a peak in the number of Risk tweets in November in 2015.

The chart in Figure 2 also displays an increase of benefit-related tweets during the month of November in 2015. A close inspection of the tweets leads us to believe that the general increase of tweets was perhaps due to the International Driverless Cars Conference that occurs annually in November.

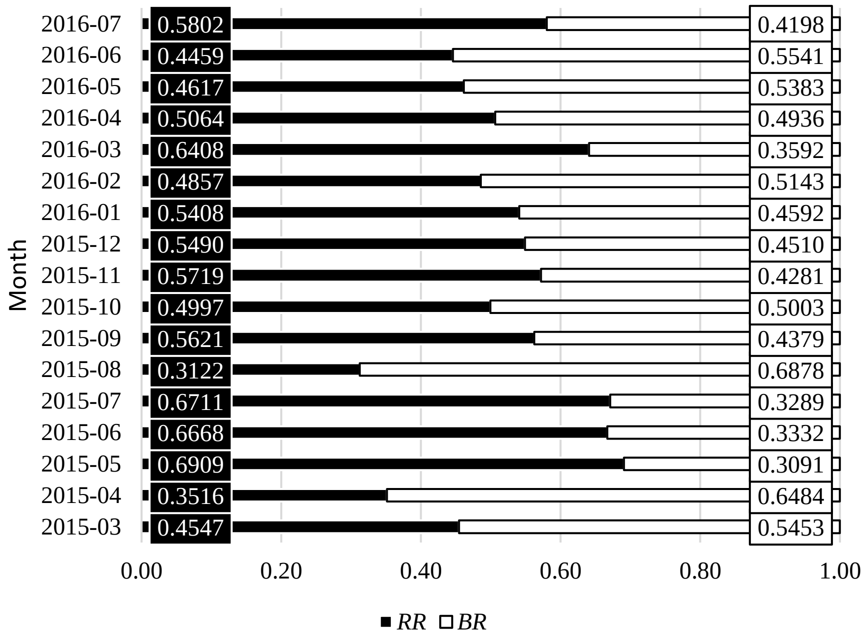


Figure 1. *RR* and *BR* over time

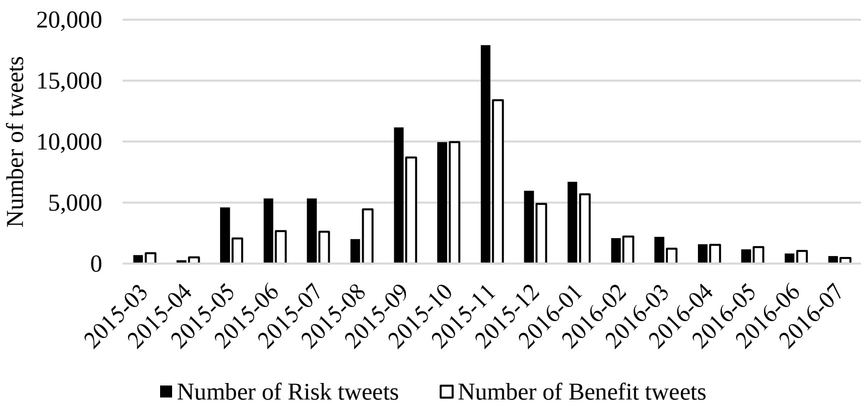


Figure 2. Number of Risk and Benefit tweets per month

5 Discussion

Before discussing the results in more detail, we discuss the limitations of this research. The tweets returned by the Twitter Search API are determined by proprietary algorithms and are not a representative sample of the overall tweets [44]. Furthermore, Twitter users are not a representative sample of the population [44]. As our analysis is based on English tweets, the main population analyzed might be located in the U.S. and should not be considered a representative global or U.S. sample [45]. In addition, only a small fraction of tweets in our dataset contain a geolocation so that we could not differentiate between different regions, which remains an ongoing issue in Twitter research [46]. As the training dataset plays an important role for training and evaluating the SVM, results depend on its quality. As described in the previous section, we found indications that the training data is of high quality but further robustness checks could provide additional evidence for the quality of the training data. Considering these limitations, however, we found interesting results that we carefully discuss in this section. This allows us the get valuable insights about people's perceptions as previous Twitter research has [14].

The RR and BR values calculated in this study indicate that people have reservations regarding self-driving cars. People tweet about risks of self-driving cars almost three times as much than about the benefits. Even if the difference might not be as big as this number suggests due to the limitations of our analysis, technology acceptance would not be guaranteed in the current state, making a disruption of individual mobility seem unlikely in the near future. This presents a problem that needs to be tackled before self-driving are sold to the public. We calculated the BR and RR values of separate years, to analyze the tweets over time and could find a small increase in RR from 0.5581 to 0.5290 (+5.2%). This might indicate that the impressive recent technical developments do not affect risk and benefit perceptions much and communication strategies should be reconsidered.

Suggestions for improvement can be derived by going over the tweet contents of the classified tweets, and trying to understand the reasons behind both risk and benefit perceptions towards self-driving cars. Among the different risk-related tweets, most of the tweets displayed concern towards the vehicles' accident, for example: "[…] Google's driverless cars have been involved in four car accidents" or "CAR CRASH Google Self Driving Cars to Decide if You Live or Die […]". This might be a case of a distorted perception of a risk as it contradicts current research. Experts argue that 93% of car accidents are due to driver error [23] and the use of self-driving cars could reduce car accidents by that exact amount [6].

People also display distrust towards the manufacturing companies and conveyed their love for driving, for example: "Sorry @google not going to buy a self driving car I like driving and don't trust your technology". In this case, benefit perception might be distorted. While driving can be enjoyable in certain situations, we find ourselves often confronted with less enjoyable aspects of driving such as traffic congestions, long monotonous highways with speed limitations, or on the search for a parking space in increasingly crowded cities. The author of this tweet might not be aware of this perspective, which could be used in communication strategies to improve benefit perceptions.

Furthermore, people also displayed fear for their own safety and privacy (e.g., "[…] Can #driverless #cars be made safe from hackers?"), where hacking someone's car could allow others to take control of your vehicle. Hackers might even go as far as writing viruses that could be transmitted from car to car. This is a risk that could proof to be real. We already see hacker attacks on current cars. These hacking attacks could cause physical harm to the passengers, which might be perceived more severe than having a personal computer hacked even if the consequences can be severe, too (e.g., huge financial losses, loss of private documents, publication of sensitive data). Manufacturers of self-driving cars need to be aware of that and provide strategies of how to avoid hacking of their vehicles.

Regarding the tweets that were classified as benefits of self-driving cars, many users were especially attracted to the fact that they could save time through self-driving vehicles, for example: “Sleepy time in the car for a in back seat. Wish I had a self driving car & I coulda joined em……”. This is also might be a case of distorted benefit perception since only full self-driving automation or level 4 automation [21] allows sleeping while driving. The current level of automation is 2 and it is likely to take some years until we arrive at level 3 or even level 4 automation. Meanwhile, many drivers are misusing current self-driving, for example by even leaving the driver’s seat entirely while driving on a public road using the Autopilot feature of a Tesla Model S [20]. People expecting to soon be able to sleep while driving might become disappointed if such systems will not be released soon as suggested by some developers of self-driving cars.

In general, people are impressed by the innovation put into the self-driving concept, for example: “[…] That hyper-futuristic driverless Mercedes has been spotted in San Fran – again […]”. Most benefit tweets reflected that people were simply excited to try something new, for example: “[…] A perk of living near Google… We saw the self-driving car today on the highway!” Developers of self-driving have recognized that people are excited about this new technology and the benefits it could provide. Consequently, they are investing in the development of self-driving cars and already promise features that will first be implemented in several years. If communication strategies are not adjusted, this excitement could cause exaggerated risk perceptions and a misunderstanding of the benefits self-driving cars are going to provide. Focusing only on the benefits and even generating exaggerated benefit perceptions could have adverse effects on public acceptance of self-driving cars.

6 Conclusion

The results indicate the need for developers and manufacturers to listen to the voice of customers of self-driving cars and probably rethink their communication strategy. By analyzing 601,778 tweets using supervised machine learning classification, we identified the need to clearly reassure the public of their risk perceptions. People tweet more about risks of self-driving cars than about the benefits. Many of the supportive tweets indicated that the benefit perceptions neglect the actual state of the technology and, thus, could be dangerous or lead to disappointment when trying the new technology for the first time. Getting potential customers to perceive the objective benefits of self-driving cars such as increased safety and increased comfort might increase benefit perception sustainably. This would lead to less disappointment with self-driving cars when they become available to the broad public and, thus, lead to higher acceptance. It is not likely that self-driving cars will disrupt individual mobility in the near future due to the lack of acceptance.

This analysis focused only on Twitter. Further research could replicate this approach using different machine learning algorithms, datasets, and other new technologies. It was not in the scope of this paper to optimize the machine learning text classification to reach the best possible classification accuracy of the SVM. By tuning the parameters of the SVM or generating additional training data, analyses could be improved. Further research of self-driving cars could be based on other keywords and use other approaches such as topic modeling [47] instead of supervised machine learning to remove the effortful manual classification of Tweets.

With the applied optimizations for text classification we could achieve sufficient accuracy of the text classification. Combined with manual inspection of the classified tweets to identify the cause for certain developments of risk and benefit perceptions, we could make well-founded suggestions for improving the public acceptance of self-driving cars. We identified a promising metric, risk rate *RR*, which can be used to study risk and benefit perceptions in social media. Furthermore, we identified issues in the communication strategies of self-driving car developers.

7 Acknowledgements

This work was performed within the Munich Center for Technology in Society (MCTS) Post/Doc Lab "Automation & Society: The Case of Highly Automated Driving".

References

1. Bower, J.L., Christensen, C.M.: Disruptive Technologies: Catching the Wave. Harv. Bus. Rev. 73, 43–53 (1995).
2. Cherubini, S., Iasevoli, G., Michelini, L.: Product-Service Systems in the Electric Car Industry: Critical Success Factors in Marketing. J. Clean. Prod. 97, 40–49 (2015).
3. Teubner, T., Flath, C.M.: The Economics of Multi-Hop Ride Sharing: Creating New Mobility Networks Through IS. Bus. Inf. Syst. Eng. 57, 311–324 (2015).
4. The Economist: If Autonomous Vehicles Rule The World: From Horseless to Driverless, http://worldif.economist.com/article/12123/horseless-driverless (Accessed: 31.10.2016).
5. Fagnant, D.J., Kockelman, K.: Preparing a Nation for Autonomous Vehicles: Opportunities, Barriers and Policy Recommendations. Transp. Res. Part A Policy Pract. 77, 167–181 (2015).
6. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Manage. Sci. 35, 982–1003 (1989).
7. Gogoll, J., Müller, J.F.: Autonomous Cars: In Favor of a Mandatory Ethics Setting. Sci. Eng. Ethics. 1–20 (2016).
8. Gupta, N., Fischer, A.R.H., Frewer, L.J.: Socio-Psychological Determinants of Public Acceptance of Technologies: A Review. Public Underst. Sci. 21, 782–95 (2012).
9. Siegrist, M.: The Influence of Trust and Perceptions of Risks and Benefits on the Acceptance of Gene Technology. Risk Anal. 20, 195–204 (2000).

10. O'Donnell, J.: Self-Driving Cars Could Have Long Road to Acceptance, http://www.usatoday.com/story/money/cars/2013/06/13/autonomous-vehicles-audi-google-drivers-crashes/2415909/ (Accessed: 31.10.2016).
11. Pizano, G.: How Much Has Social Media Changed Society?, http://www.shoutmeloud.com/how-much-has-social-media-changed-society.html (Accessed: 31.10.2016).
12. St Louis, C., Zorlu, G.: Can Twitter Predict Disease Outbreaks? BMJ. 344, e2353 (2012).
13. Tumasjan, A., Sprenger, T.O., Sandner, P.G., Welpe, I.M.: Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. In: Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media. pp. 178–185 (2010).
14. Pak, A., Paroubek, P.: Twitter as a Corpus for Sentiment Analysis and Opinion Mining. In: Proceedings of the Seventh Conference on International Language Resources and Evaluation. pp. 1320–1326 (2010).
15. Venkatesh, V., Davis, F.D., Morris, M.G.: Dead Or Alive? The Development, Trajectory And Future Of Technology Adoption Research. J. Assoc. Inf. Syst. 8, 267–286 (2007).
16. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User Acceptance of Information Technology: Toward a Unified View. Source MIS Q. 27, 425–478 (2003).
17. Venkatesh, V., Thong, J.Y.L., Xu, X.: Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead. J. Assoc. Inf. Syst. 17, 328–376 (2016).
18. Martins, C., Oliveira, T., Popovič, A.: Understanding the Internet Banking Adoption: A Unified Theory of Acceptance and Use of Technology and Perceived Risk Application. Int. J. Inf. Manage. 34, 1–13 (2014).
19. Lancelot Miltgen, C., Popovič, A., Oliveira, T.: Determinants of End-User Acceptance of Biometrics: Integrating the "Big 3" of Technology Acceptance with Privacy Context. Decis. Support Syst. 56, 103–114 (2013).
20. Yadron, D., Tynan, D.: Tesla Driver Dies in First Fatal Crash While Using Autopilot Mode, https://www.theguardian.com/technology/2016/jun/30/tesla-autopilot-death-self-driving-car-elon-musk (Accessed: 31.10.2016).
21. NHTSA: National Highway Traffic Safety Administration Preliminary Statement of Policy Concerning Automated Vehicles, http://www.nhtsa.gov/About+NHTSA/Press+Releases/U.S.+Department+of+Transportation+Releases+Policy+on+Automated+Vehicle+Development (Accessed: 31.10.2016).
22. Krok, A.: This Is the Stupidest Misuse of Tesla's Autopilot yet, http://www.cnet.com/news/this-is-the-stupidest-misuse-of-teslas-autopilot-yet/ (Accessed: 31.10.2016).
23. Gold, C., Dambock, D., Lorenz, L., Bengler, K.: "Take over!" How Long Does It Take to Get the Driver Back into the Loop? In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. pp. 1938–1942 (2013).
24. Schoettle, B., Sivak, M.: A Survey of Public Opinion About Autonomous and Self-Driving Vehicles in the U.S., the U.K., and Australia, https://deepblue.lib.umich.edu/handle/2027.42/108384 (Accessed: 31.10.2016).
25. Treat, J.R., Tumbas, N.S., McDonald, S.T., Shinar, D., Hume, R.D., Mayer, R.E., Stansifer, R.L., Castellan, N.J.: Tri-Level Study of the Causes of Traffic Accidents: Final Report, https://trid.trb.org/view.aspx?id=144150 (Accessed: 31.10.2016).
26. Fraedrich, E., Cyganski, R., Wolf, I., Lenz, B.: User Perspectives on Autonomous Driving. In: Arbeitsberichte 187. Geographisches Institut, Humboldt-Universität, Berlin (2016).
27. Eckoldt, K., Knobel, M., Hassenzahl, M., Schumann, J.: An Experiential Perspective on Advanced Driver Assistance Systems. it - Inf. Technol. 54, 165–171 (2012).

28. Kasperson, R.E., Kasperson, J.X.: The Social Amplification and Attenuation of Risk. Ann. Am. Acad. Pol. Soc. Sci. 545, 95–105 (1996).
29. Kleijnen, M., Lee, N., Wetzels, M.: An Exploration of Consumer Resistance to Innovation and Its Antecedents. J. Econ. Psychol. 30, 344–357 (2009).
30. Pang, B., Lee, L.: Opinion Mining and Sentiment Analysis. Found. Trends Inf. Retr. 2, 1–135 (2008).
31. Sharma, R., Yetton, P., Crawford, J.: Estimating the Effect of Common Method Variance: The Method-Method Pair Technique with an Illustration from TAM Research. MIS Q. 33, 473–490 (2009).
32. Okazaki, S., Diaz-Martin, A.M., Rozano, M., Menendez-Benito, H.: How to Mine Brand Tweets Procedural Guidelines and Pretest. Int. J. Mark. Res. 56, 467–489 (2014).
33. Twitter: The Search API, https://dev.twitter.com/rest/public/search (Accessed: 31.10.2016).
34. Twitter: Public API: GET Search/tweets, https://dev.twitter.com/rest/reference/get/search/tweets (Accessed: 31.10.2016).
35. Twitter: Help Center: The Basics, https://support.twitter.com/articles/131209 (Accessed: 31.10.2016).
36. Feinerer, I., Hornik, K., Meyer, D.: Text Mining Infrastructure in R. J. Stat. Softw. 25, 1–54 (2008).
37. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? Sentiment Classification Using Machine Learning Techniques. In: Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - EMNLP ’02. pp. 79–86 (2002).
38. Joachims, T.: Text Categorization with Support Vector Machines: Learning with Many Relevant Features. In: Nédellec, C. and Rouveirol, C. (eds.) Machine Learning: ECML-98. pp. 137–142. Springer Berlin Heidelberg (1998).
39. Chang, C., Lin, C.: LIBSVM : A Library for Support Vector Machines. ACM Trans. Intell. Syst. Technol. 2, 1–39 (2011).
40. Kohavi, R.: A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. Int. Jt. Conf. Artif. Intell. 14, 1137–1143 (1995).
41. Sriram, B., Fuhry, D., Demir, E., Ferhatosmanoglu, H., Demirbas, M.: Short Text Classification in Twitter to Improve Information Filtering. In: Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR ’10. pp. 841–842 (2010).
42. Kuhn, M.: Building Predictive Models in R Using the Caret Package. J. Stat. Softw. 28, 1–26 (2008).
43. Rubinkam, M.: Driverless Truck Meant to Improve Safety in Work Zones, https://www.yahoo.com/news/driverless-truck-meant-improve-safety-zones-202055180.html (Accessed: 31.10.2016).
44. Ruths, D., Pfeffer, J.: Social Media for Large Studies of Behavior. Science. 346, 1063–1064 (2014).
45. Mislove, A., Lehmann, S., Ahn, Y.-Y., Onnela, J.-P., Rosenquist, J.N.: Understanding the Demographics of Twitter Users. In: ICWSM ’11 Proceedings of international AAAI Conference on Weblogs and Social Media. pp. 554–557 (2011).
46. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, K.P.: Measuring User Influence in Twitter: The Million Follower Fallacy. In: ICWSM ’10: Proceedings of international AAAI Conference on Weblogs and Social Media. pp. 10–17 (2010).
47. Debortoli, S., Müller, O., Junglas, I., Vom Brocke, J.: Communications of the Association for Information Systems Text Mining For Information Systems Researchers: An Annotated Topic Modeling Tutorial. Commun. Assoc. Inf. Syst. 39, 110–135 (2016).

Smart Meter Data Analytics for Enhanced Energy Efficiency in the Residential Sector

Mariya Sodenkamp[1], Ilya Kozlovskiy[1], Konstantin Hopf[1], Thorsten Staake[12]

[1]Energy Efficient Systems Group, University of Bamberg
[2]Department of Management, Technology and Economics, ETH Zurich
{mariya.sodenkamp, ilya.kozlovskiy, konstantin.hopf, thorsten.staake}@uni-bamberg.de

Abstract. Achievement of the ambitious environmental sustainability targets requires improvement of energy efficiency practices in private households. We demonstrate how utility companies, having access to smart electricity meter data, can automatically extract household characteristics related to energy efficiency and adoption of renewable energy technologies (e.g., water/space heating type, age of house, number and age of electric appliances, interest in installation of photovoltaic systems etc.) by using supervised-machine-learning-based green IT artifacts. The gained information enables design of custom-tailored interventions (such as promotion of personalized energy audits, ecologic services and products, or load shifting mechanisms) that trigger residents' behavioral change toward environmental sustainability as well as improvement of utilities' key performance indicators. Moreover, realizing privacy preservation concerns, we investigate the influence of smart meter data granularity and the amount of survey responses required for the artifact development on the household classification quality.

Keywords: Green information systems (IS), smart meters, data analytics, energy efficiency, sustainability

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Sodenkamp, M.; Kozlovskiy, I.; Hopf, K.; Staake, T. (2017): Smart Meter Data Analytics for Enhanced Energy Efficiency in the Residential Sector, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1235-1249

1 Introduction

In order to promote energy efficiency and the integration of renewable energy sources, policy makers have placed high hopes on networked electricity meters that measure and communicate consumption information at a high resolution in time [1]. These so-called smart meters offer, besides improvements to the utilities' billing processes, timely consumption feedback for residential customers, render dynamic tariff schemas possible, and provide input to home automation systems. Ultimately, smart meters should help citizens to reduce their electricity consumption and motivate shifting loads to support the integration of renewable, fluctuating electricity sources. Early studies have reported promising effects including energy savings as high as 15% [2], and many countries mandated utility companies to roll out the technology among their residential customers. To date, more than 50 million smart meters have been deployed in the EU, and industry expects 154 million devices by 2017 [3].

The initial enthusiasm for the technology, however, has faded recently. More carefully designed studies that were not affected by sampling bias (e.g., from primarily observing volunteering, motivated users) showed savings of only around 3% and little response to incentives for load shifting [4]. Moreover, the mostly poor design of feedback campaigns together with deficient data protection practices raised substantial privacy concerns among consumers [5]. It has become evident that relying on the mere effects of consumption information and dynamic tariffs is not sufficient to advance energy literacy, motivate the hoped-for behavioral change, nor to trigger investments in saving technologies [6].

We argue that the current, disappointing performance of smart metering is not a problem of the technology as such, but resides in an insufficient information extraction from the available metering data. Feedback interventions, for example, that are tailored to individual recipients have been shown to achieve substantially higher saving effects and to reach a better user acceptance [7]. Such measures (e.g., providing concrete advice or comparisons to similar households, offering services that reflect the household's characteristics, etc.) typically yield considerably lower cost per kWh saved (or shifted) than tax credits and rebates and meet higher public acceptance than prohibitive regulations [8]. Load shifting measures, control of heating systems or combined micro heat and power plants can also considerably benefit from information beyond pure metering data if, for example, up-to-date knowledge on the presence of the inhabitants is available. While strong evidence supports the benefits of using such specific information on households to conduct consumer-specific energy efficiency campaigns [8], a major problem is, that the required data to conduct such campaigns is not available for large scale deployments. This reduces the benefits of smart metering infrastructures dramatically.

The research outlined herein lies within the scope of energy informatics - an emerging discipline concerned with analysis, design and implementation of information systems to reduce energy consumption [9, 10]. ***We strive for providing the missing link between smart meter data and powerful energy efficiency measures***. We show how data collected at 15-minute granularity from out-of-the-shelf smart electricity meters can be used to infer energy-efficiency related characteristics of residential dwellings (e.g., water/space heating type, age of house, number and age of electric appliances, interest on installation of photovoltaic systems etc.), ***using supervised machine learning*** techniques. Automatically mining this information – with the consent of inhabitants – enables large-scale, targeted saving advice and better input to home automation systems that can significantly contribute to reduction of energy consumption and advancement of environmental sustainability.

2 Research Objective and Theoretical Context

A number of researchers have developed methods to predict household characteristics data from energy consumption data. The approaches differ with respect to the type of data available (e.g., load records, household survey reports), their resolution in time, and the output variables of interest. The potential to recognize characteristics from load curves depends on the data granularity – the finer the granularity is the better recognition performance can be achieved. Vast research has been done on the recognition of devices from the data of extreme granularity (Hertz and Megahertz frequency) – in the field of Non Intrusive Load Monitoring (NILM) [11, 12]. While sampling rates far beyond one megahertz are quite common in industrial or lab settings, the smart metering infrastructure that is currently being deployed and is expected to be in the field for the next 20 years does not provide such fine grained data. Therefore, the NILM methods will not be compatible with the standard meters deployed in most households [13].

Several authors have investigated coarser consumption data. Chicco [14] provides an overview of the clustering methods for electrical load pattern grouping. In the field of recognition of energy efficiency characteristics which we focus on, Fei et al.[15], proposed a method to detect heat pumps from the daily energy consumption data. Beckel et al. [16] used 30-minute data reduced to 26 features to infer 18 household properties, most of which relate to the inhabitants' life situation (age, family, employment, social class, etc.), and three directly relate to energy efficiency (number of appliances, cooking type and lightbulbs). Hopf et al. [17] and Sodenkamp et al. [18] improved Beckel's algorithm by extracting 88 features, applying filtering methods and by refining the properties. Sodenkmp et al. [19] included weather data to Beckel's algorithm using a multi-dimensional classification method called DID-class.

Further works employ conventional yearly consumption readings for dwelling classification. Kozlovskiy et al. [20] detected old gas heating systems for a targeted cross-selling campaign in Belgium. Sodenkamp et al. [21] predicted household probability to register on an energy efficiency portal for customer engagement campaigns in Germany and Switzerland. Hopf et al. [22] used yearly consumption and geographic data from OpenStreetMap and GeoNames to detect living area, household type and number of residents.

Our overarching goal is to **examine the usage of machine-learning-based smart meter data analytics methods in the practice of energy utility companies and their contribution to the environmental and economic sustainability.** In this work, we go beyond the state-of-the art by identifying eleven energy-efficiency related household characteristics (space heating type and age, water heating type, heat pump usage in a household, house age, number and age of appliances, presence and interest in photovoltaic and thermal installations, number of recently completed energy efficiency measures, and type of cooking facility) from smart meter data at 15-minute granularity, which we collected from a Swiss utility company. We extract 93 features from the consumption data and include weather data represented by 40 features. Finally, we investigate the effects of data granularity. This information can be used for the development of targeted energy-efficiency measures (e.g., saving tips, promotions of installations of renewable energy systems, load shifting campaigns). To enable algorithm training and testing we designed and conducted an online survey. Thus, our first research question is as follows:

RQ 1. *To what extent is it possible to recognize energy efficiency related household characteristics from smart meter data using machine learning methods?*

The availability of detailed electricity consumption data-traces to the utilities is associated with consumers' privacy concerns [5]. Depending on the company policy and local legislation, utilities make use of different data granularities. The typical aggregations vary between 15-min, 30-min, hourly and daily levels. For daily data, some utilities differentiate between the HT (high tariff, during the day) and NT (low tariff, during the night) consumption. Data with lower frequency contains less information about customers' behavior and we expect the performance of our classification algorithms to degrade when applied to such data. Therefore, we test how different data granularities influence the resulting classification performance and formulate our second research question as follows:

RQ 2*. To what extent do different granularities of smart meter data influence the recognition quality of energy efficiency related household characteristics?*

3 Data Description

For our data science study, we cooperated with a utility company in Switzerland with about 9`000 customers that provided us with household electricity meter readings at 15-minute granularity in the timespan between June 1st 2014 and May 31st 2015. For the same time period, we acquired hourly weather-data from the U.S. National Climate Data Center [23]. Together with the utility, we conducted a web-based customer survey about energy efficiency related household characteristics between June and September 2015. All customers received an invitation to the survey attached to their bimonthly bill. In this survey, we collected data on 527 households, which corresponds to a response rate of 6%. We matched the survey results with smart meter data using respondents' names and addresses.

Based on the survey responses, we defined and extracted 11 energy efficiency relevant household characteristics (*properties*) that include at least two *classes* (see Table 1). The class definition was either naturally given (e.g., heat pump exists / does not exist) or we empirically set class borders by using quantiles. When we used quantiles for defining the class border, we either aimed to separate the households in equally sized classes, or wanted to identify a specially interesting class (e.g., high purchasing intention for solar installation). The total number of households per property does not necessarily respond to the total number of households participating in the survey, since the survey participants left some questions unanswered. The amount of excluded instances due to the missing data is different per property and lies mostly at 23.48 %, except for "Age of residency" with 23.91%, "Age of appliances" with 34.35%, and "Age of heating" with 40% of all survey participants excluded.

Table 1: Energy efficiency related household properties with classes defined from survey responses; q_x denotes statistical x percentile of the survey responses (relative class sizes are not necessarily identical to the percentiles by definition for numerical values, due to categorical survey variables).

Property	**Classes**	**Definition**	**Class size**	
			Abs.	**Rel.**
Age of appliances	New	Avg. appliance age < $q_{0,25}$	153	33,41%
	Average	Avg. appliance age between $q_{0,25}$ and $q_{0,75}$	152	33,19%
	Old	Avg. appliance age > $q_{0,75}$	153	33,41%
Num. of appliances	Few	Number appliances < $q_{0,25}$	149	28,27%
	Average	Number appliances between $q_{0,25}$ and $q_{0,75}$	280	53,13%
	Many	Number appliances > $q_{0,75}$	98	18,60%
Cooking type	Electric	Number electric stoves > 0	484	91,84%
	Not electric	Number electric stoves = 0	43	8,16%
Efficiency measure	No	Number completed energy efficiency measures during the last 15 years (insulation of basement / roof / building envelop, or window replacement)	304	57,69%
	Few		109	20,68%
	Multiple		114	21,63%
Heat pump	No	Existing heat pump	453	85,96%
	Yes		74	14,04%
Age of residency	< 10	Age (in years) of the building the household is living in	71	13,65%
	10-29		147	28,27%
	30-74		219	42,12%
	≥ 75		83	15,96%
Interest in solar	Low	Purchase intention coefficient < $q_{0,50}$	387	73,43%
	Average	Purchase intention coefficient between $q_{0,50}$ and $q_{0,75}$	49	9,30%
	High	Purchase intention coefficient > $q_{0,75}$	91	17,27%
Solar installation	Yes	Photovoltaics or solar heating existent	29	5,50%
	No	Neither photovoltaics nor solar heating existent	498	94,50%
Age of heating	New	Space heating age < $q_{1/3}$	128	33,51%
	Average	Space heating age between $q_{1/3}$ and $q_{2/3}$	135	35,34%
	Old	Space heating age > $q_{2/3}$	119	31,15%
Space heating type	Electric	Space heating = „*Electric heating*"	21	3,98%
	Heat pump	Space heating = „*Heat pump*"	66	12,52%
	Other	Other space heating	440	83,49%
Water heating type	Electric	Water heating = „*Electric heating*"	81	15,37%
	Heat pump	Water heating = „*Heat pump*"	63	11,95%
	Other	Other water heating	383	72,68%

4 Analysis

In this section, we describe our data analysis methodology and results. We first seek to answer the first research question:

RQ 1. *To what extent is it possible to recognize energy efficiency related household characteristics from smart meter data using machine learning methods?*

To answer this question, we follow the four-step procedure described below.

Step 1: Feature Extraction. At the beginning, it is important to reduce the raw data dimensionality and transform the data to a more usable form. Consumption time series are divided into single weeks, since from previous work we know that a weekly energy consumption is sufficient to perform household classification [16]. We extracted *93 features from 15-min smart meter data* for each week, that are adopted from previous works dealing with 30-min smart meter data [16, 17]. The details can be found in the project report [24]. The extracted features cover four categories: consumption (e.g., in the morning, noon, evening); ratios of consumption figures (e.g., consumption in the morning vs. noon, daytime vs. night); statistics (e.g., variance, quantiles); others (e.g., number and average heights of consumption peaks).

Besides the smart meter features, we defined and used *40 features describing the correlation between electricity consumption and weather data*, since a positive effect of weather data on the classification performance was shown in the previous study [19]. For each weather variable (temperature, wind speed, sky cover, and precipitation), we calculate eight features: overall correlation over the week, correlation during the day and during times of the day (night, daytime, evening), correlation of minima in both time series, correlation of weather minima and consumption maxima, and ratio of the weekday and weekend correlations.

Finally, we use 133 features for our analysis. Due to the space constraints we cannot present all the features (interested reader is referred to [24]), but we list 10 most frequently selected features in the final prediction models for all household properties with a short description in Table 2.

Step 2: Feature selection. After having prepared the feature vectors, we select relevant features for each property separately. This is done to reduce overfitting and speed up the calculations by removing the irrelevant features. We tested the following feature selection methods: Correlation based (cfs), consistency based (consistency), Based on the importance from random forest (importance), statistical test for the difference in distributions (chi.squared), entropy based (gain.ratio), forward feature selection search (forward-selection), backward feature selection search (backward-selection), no feature selection (none). For the methods description see [25].

Table 2: The top 10 selected features and their description

Rank	Feature	Description
1	t15_above_2kw	Time with consumption above 2 kW
2	t15_value_min_guess	Time with consumption above minimal consumption
3	r15_wd_evening_noon	Relation between evening and noon consumption on weekdays
4	r15_mean_max_no_min	Relation between mean and max consumption with subtracted

		minimum
5	t15_time_above_base2	Time with consumption above estimated baseline
6	r15_evening_noon	Relation between evening and noon consumption
7	t15_const_time	Time with nearly constant consumption
8	t15_daily_min	Average daily minimum
9	r15_min_wd_we	Relation between minimum of consumption on weekdays and weekends
10	r15_var_wd_we	Relation between the variance of consumption on weekdays and weekends

Step 3: Classification. As the next step we train the model based on the selected features and evaluate the following six well-known classifiers that have implementations in the statistical programming environment GNU R: AdaBoost [26], k Nearest Neighbors (kNN) [27], Naïve Bayes [27], Random Forest [28], Support Vector Machine (SVM) [29].

Step 4: Evaluation. To measure the classification performance, we are interested in testing how many households are correctly classified in each available class. These numbers are typically presented in the form of a confusion matrix. To compare the different confusion matrixes between each other we calculate two measures from the confusion matrices:

- Accuracy**:** It is defined as the portion of correctly classified instances from the number of total classification instances and can take values between 0 and 1, where 1 corresponds to perfect prediction and 0 to total misclassification. Accuracy is easy to interpret, but in the situation where the classes are unbalanced (i.e., one class occurs much more often than the others) a classifier that always predicts a majority class can achieve high accuracy. Therefore, this measure can be slightly misleading if applied to such unbalanced properties.
- Matthews Correlation Coefficient (MCC): It is an alternative measure that is more suitable for the unbalanced problems. It is a correlation coefficient between the observed and predicted classifications. In the case of binary classification problem, it is equal with the phi statistic [30]. We use MCC definition for multiclass problems [31, 32]. MCC can take values between -1 and 1, where 1 corresponds to the perfect classification, -1 to the total disagreement between the predictions and real observations and 0 for the classification that is not better than random prediction. MCC lacks the easy interpretability of the accuracy measure, but it is a good compromise among discriminancy, consistency and coherent behaviors with varying number of classes, unbalanced datasets and randomization [31].

To calculate the performance measures, we first split the data into the *test set* (10%) from the *main data* by using a stratified split (the distribution of classes in the test set deviates at most with one household from the distribution in the main data). For the main data, we use 4-fold cross-validation to select the best classification algorithm and feature selection method.

Then, we take the features describing a single week (week number 34, from 12.01 to 18.01.2015). The features of the main dataset are centered and scaled to have mean

0 and standard deviation 1. The features in the test set are then scaled with the same proportions. All households with missing or not available values for the chosen week are removed, as well as the feature vectors that are constant for all households. This problem occurs mainly with weather features that are constant during the week (e.g., precipitation and sky cover).

Performance of the individual classifiers for week 34 is presented in Figure 1, together with two benchmark measures: random guess ($RG = 1/K$, where K is the number of classes in one property) and biased random guess ($BRG = \sum h^2$, where h is the relative class size of each class for one property, as displayed in Table 1). The figure illustrates, that no single classifier provides the best classification performance for all properties. Similarly, no best feature selection technique for all properties can be found. Therefore, we choose the classification configuration (feature selection and classifier) that produced the best result for MCC and list it in Table 3.

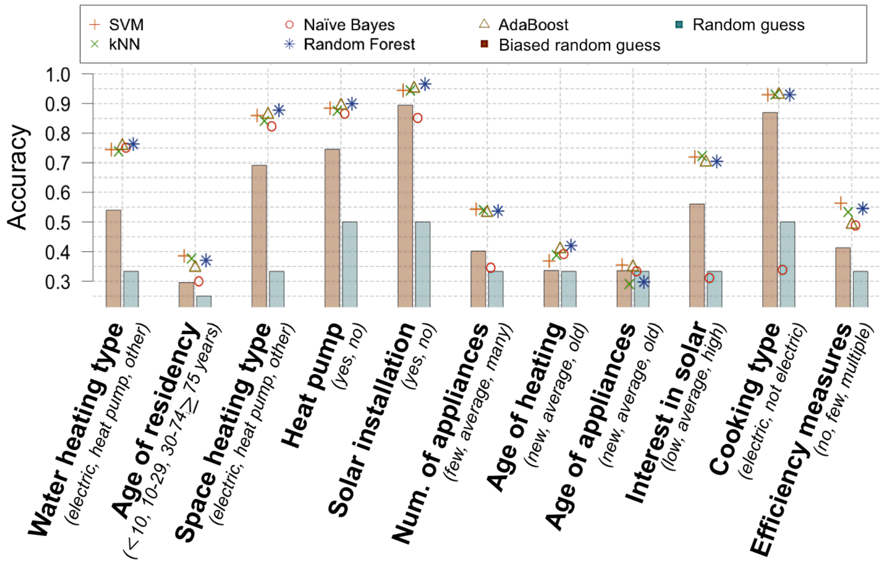


Figure 1: Comparison of classification performance of individual classifiers without feature selection (week number 34)

With these best performing classifier configurations, the final model is trained on all households that are not in the test set. Predictions are then made for the test set and are then evaluated in terms of both the accuracy and MCC. We repeated the training and prediction process for all 29 week of data that were not excluded due to public or school holidays or missing values. The results for every week are then aggregated with a simple ensemble classifier [33], that averages the prediction probabilities for each class and thus forms one single multi-week-classifier. The final predictions for the test data are then evaluated with respect to both the accuracy and MCC. The results are presented in the Figures 2 and 3.

As the result, we can answer our first research question positively: It is possible to predict most (9 from 11) energy efficiency relevant household properties better than by using random guess. We especially achieved good results for the prediction of properties related to the room and water heating (*Space heating type*, *Water heating type*, and *Heat pump*). The properties *Cooking type* and *Efficiency measures* could not be predicted adequately with our approach. But even with the negative MCC, the results can still be valuable for utility companies that usually do not know what household belong to which class: the application of the classifier can perform better than random guessing.

Table 3: The best performing configurations

Property	Feature selection method	Classifier
Space heating type	cfs	Random Forest
Water heating type	none	Random Forest
Heating age	gain.ratio	Support Vector Machine
Age of residency	cfs	AdaBoost
Age of appliances	chi.squared	Naïve Bayes
Cooking type	none	Random Forest
Heat pump	chi.squared	Random Forest
Solar installation	none	Random Forest
Efficiency measures	cfs	AdaBoost
Num. of appliances	none	Random Forest
Interest in solar	chi.squared	Random Forest

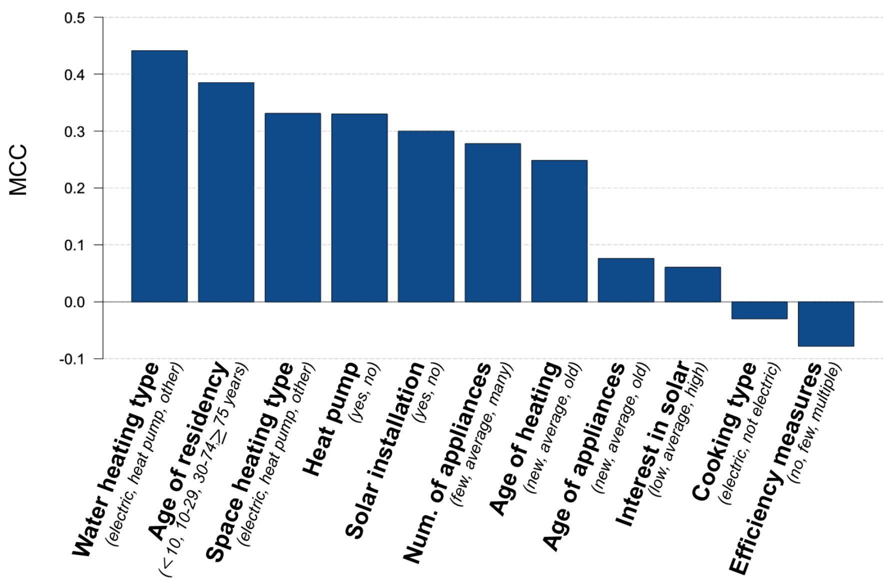


Figure 2: Classification performance for all properties measured with MCC

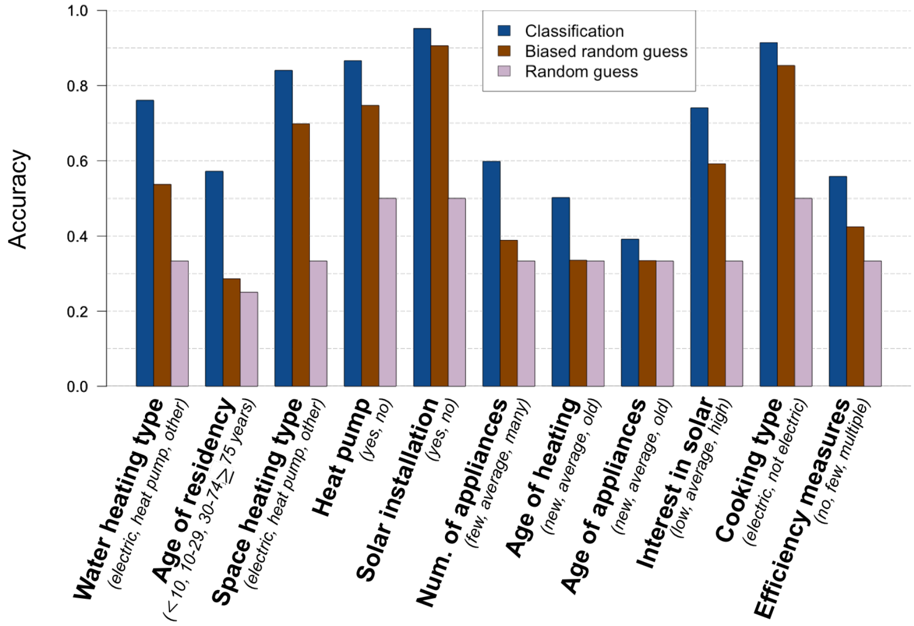


Figure 3: Classification results for all properties measured with accuracy, compared to random guess and biased random guess

RQ 2. *To what extent do different granularities of smart meter data influence the recognition quality of energy efficiency related household characteristics?*

To answer this question, we have to repeat the classification process as described in RQ1 for different data granularities. We simulate the different granularities by aggregating the existing 15-minute smart meter data up to the following typical aggregation levels: 15-minute, 30-minute, 60-minute, daily NT/HT (low-tariff during the night / high-tariff during the day), and daily. The NT consumption is measured during different times by different utility companies. For this case we will assume that NT consumption is measured between 23:00 and 07:00.

For each data granularity, we adapted the defined features that are reasonable for the data granularity: For 30-minute and 60-minute data we use features analogously to the 15-minute ones. For the NT/HT and daily aggregation levels we define 14 and 7 different features respectively that describe the consumption during the weekdays and weekends, the relations between different consumptions and the variance. Naturally, a large set of the 15-min smart meter features cannot be calculated for HT/NT or daily measurements (e.g., features on consumption during times of the day, or peaks of the load curve). Additionally, we compare the results to the worst case of only having a single weekly value (1 feature).

The classification is repeated for a single week only. We exclude the weather variables from this analysis, because we want only to show the value of the data at different granularities. The consideration of weather will have more value for the finer granularities, since it is also available on the 15-minute level. For calculations at each granularity levels, the features from coarser granularities are also included. This is done to ensure that important features are not missed on any granularity level. E.g., if there is a good feature that we calculate based on the daily data, but do not for hourly data, then we will get better results for daily granularity than for the hourly data, even though we could get the same or better result by including this feature with hourly data. In this way, we get a large number of features, especially for the 15-minute data, and therefore we perform feature selection with the three best performing methods on the 15-min data (cfs, chi.squared, gain.ratio), and without feature selection.

In Figures 4 and 5, we show the best classification performance that was achieved for different granularities with one of three feature selection methods or without feature selection using Random Forest classifier. Since our goal in this analysis is to compare the best possible performances of different data granularities, we do not create overfitted models that are not necessarily designed to predict household classes for new households, and use the same data for training and test. Therefore, the results cannot be compared with those presented in Fig. 2 and 3, but they give an impression on the impact of data granularity on the classification performance. We can conclude from the results, that ***there is not much difference between using the 15-, 30- and 60-minute data*** for the prediction of the energy efficiency relevant household characteristics. There is a large drop in performance by using the daily (HT/NT and 24-hour) data for some properties. Using the weekly values shows the worst performance, demonstrating the value of finer granularity data.

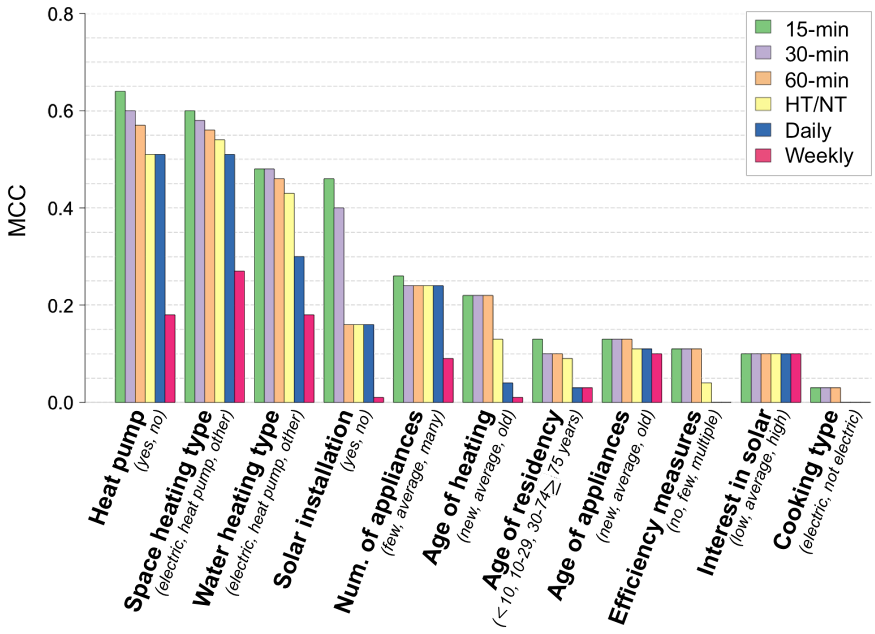


Figure 4: MCC results for different data granularities

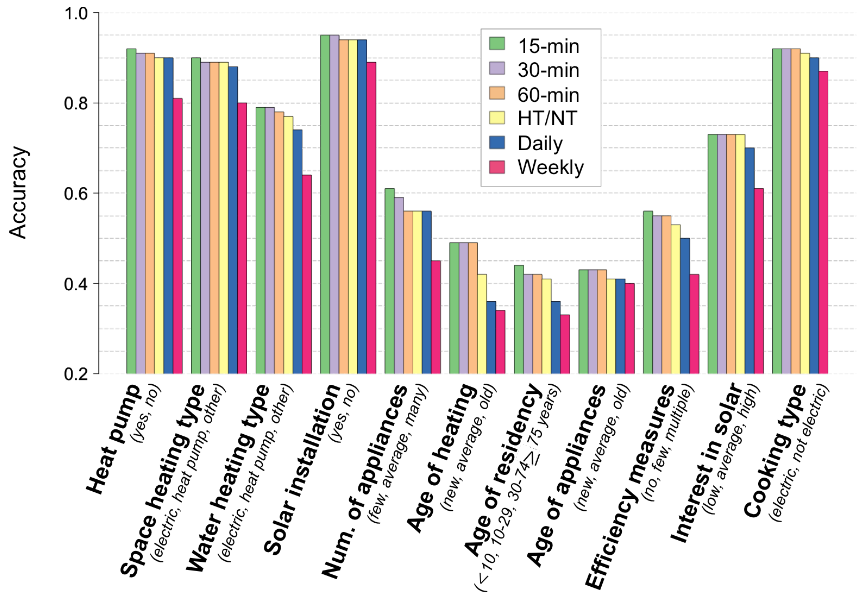


Figure 5: Accuracy results for different data granularities

5 Discussion and Conclusions

The objective of our study was the identification of energy-efficiency and renewable energy related household characteristics based on smart electricity meter data. For the development of our methodology and the training of our machine-learning based green IT artifact, we used smart meter data collected at 15-minute granularity from a Swiss utility company and conducted a customer survey in 2015. In this paper, we have shown that it is possible to identify 9 from 11 proposed energy efficiency relevant household properties (including space heating type and age, water heating type, existence of heat pumps, number and age of appliances, existence and interest in solar installations, and house age) with average accuracy over all properties of 70%. The artifact could not identify the properties "number of completed energy efficiency measures" and "type of cooking facility". Furthermore, we have shown that the smart meter data aggregated at hourly values is sufficient to detect 8 from 11 properties with only a small loss of prediction quality, compared to 15-minute data. Importantly, tools for the recognition of household characteristics must be used under clearly documented privacy preservation conditions and with user consent. Development of law-based guidelines for energy consumption data treatment in the analytics tasks is an important task that should be solved in the next future.

We have identified the following limitations of our work. First, this study is based on the online survey implying the selection bias, which means that we cannot be sure that our sample used for the algorithms training and test is representative and that the results hold for all utility customers. Further, we worked together with a utility company that serves customers from one city and the surrounding area - this introduces a regional bias into the analysis. In addition, we only considered 11 household properties in this work - the results may differ when new properties are included.

In the future work, we plan to conduct field studies to investigate the effects of personalized interventions toward the households selected using the presented algorithms. The interventions can include offers of energy-efficiency products and services, customized consultancy on energy efficiency measures, or normative feedback. We could show the economic, ecologic and social potential of such interventions in a recent study [20] where households with inefficient heating systems were identified based on annual gas consumption data. We expect much better results with the use of smart meter consumption data as shown in this paper. In this setting, we also plan to recognize combined properties (e.g., old house without conducted energy-efficiency measures, or homeowners that are interested in solar installations) from the electricity consumption and weather data. We also plan to reproduce and expand the presented results by cooperating with utility companies serving customers from other geographic locations. A more detailed approach would investigate how the combined changes in granularity and data volume affect the classification performance. Additionally, we used only a simple ensemble learner, that computes the mean value from individual predictions, to aggregate classification results from multiple weeks. Using a more advanced approach that takes the varying performance during different seasons into account could further improve the results.

Being an example for a data science research, our work still allows for empirical validation of the effects from applying the developed artifact in field. Moreover, we demonstrate how to integrate the end-users (utility customers) in the IS research. Ultimately, the proposed artifact is applicable to virtually every smart meter deployment worldwide without changes in the hardware, and thus can ***considerably contribute to the society's energy targets***.

6 Acknowledgements

The research presented herein was financially supported by the Swiss Federal Office of Energy (Grant SI/501202-01) and EUROSTARS (Grant E!9859 - BENgine II).

References

1. European Commission: Commission Recommendation of 9 March 2012 on Preparations for the Roll-out of Smart Metering Systems, http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32012H0148, (2012).
2. Darby, S.: The Effectiveness of Feedback on Energy Consumption. University of Oxford (2006).
3. Berg Insight AB: Smart Metering in Europe. , Sweden.

4. Degen, K., Efferson, C., Frei, F., Goette, L., Lalive, R.: Smart Metering, Beratung oder Sozialer Vergleich: Was beeinflusst den Elektrizitätsverbrauch? 232 (2013).

5. McKenna, E., Richardson, I., Thomson, M.: Smart Meter Data: Balancing Consumer Privacy Concerns with Legitimate Applications. Energy Policy. 41, 807–814 (2012).

6. Graml, T., Loock, C.-M., Baeriswyl, M., Staake, T.: Improving Residential Energy Consumption at Large Using Persuasive Systems. In: Proceedings of the 19th European Conference on Information Systems (ECIS). AISeL, Helsinki, Finland (2011).

7. Loock, C.-M., Staake, T., Landwehr, J.: Green IS Design and Energy Conservation: an Empirical Investigation of Social Normative Feedback. In: Proceedings of the 32. International Conference on Information Systems (ICIS). , Shanghai, China (2011).

8. Allcott, H., Mullainathan, S.: Behavior and Energy Policy. Science. 327, 1204–1205 (2010).

9. Watson, R.T., Boudreau, M.-C., Chen, A.J.: Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS community.(Essay). MIS Q. 34, 23 (2010).

10. Watson, R.T., Howells, J., Boudreau, M.-C.: Energy Informatics: Initial Thoughts on Data and Process Management. In: Brocke, J. vom, Seidel, S., and Recker, J. (eds.) Green Business Process Management. pp. 147–159. Springer Berlin Heidelberg (2012).

11. Birt, B.J., Newsham, G.R., Beausoleil-Morrison, I., Armstrong, M.M., Saldanha, N., Rowlands, I.H.: Disaggregating Categories of Electrical Energy End-Use from Whole-House Hourly Data. Energy Build. 50, 93–102 (2012).

12. Hart, G.W.: Nonintrusive Appliance Load Monitoring. Proc. IEEE. 80, 1870–1891 (1992).

13. Kim, H., Marwah, M., Arlitt, M., Lyon, G., Han, J.: Unsupervised Disaggregation of Low Frequency Power Measurements. In: Proceedings of the 2011 SIAM International Conference on Data Mining. pp. 747–758. Society for Industrial and Applied Mathematics (2011).

14. G. Chicco, Overview and Performance Assessment of the Clustering Methods for Electrical Load Pattern Grouping, Energy 42 (1) (2012) 68–80.

15. Fei, H., Kim, Y., Sahu, S., Naphade, M., Mamidipalli, S.K., Hutchinson, J.: Heat Pump Detection from Coarse Grained Smart Meter Data with Positive and Unlabeled Learning. In: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 1330–1338. ACM, New York, NY, USA (2013).

16. Beckel, C., Sadamori, L., Staake, T., Santini, S.: Revealing Household Characteristics from Smart Meter Data. Energy. 78, 397–410 (2014).

17. Hopf, K., Sodenkamp, M., Kozlovskiy, I., Staake, T.: Feature Extraction and Filtering for Household Classification based on Smart Electricity Meter Data. In: Computer Science-Research and Development. pp. 141–148. Springer Berlin Heidelberg, Zürich (2014).

18. Sodenkamp, M., Hopf, K. and Staake, T.: Using Supervised Machine Learning to Explore Energy Consumption Data in Private Sector housing. In: Handbook of Research on Organizational Transformations through Big Data Analytics, pp.320-333 (2014).

19. Sodenkamp, M., Kozlovskiy, I., Staake, T.: Supervised Classification with Interdependent Variables to Support Targeted Energy Efficiency Measures in the Residential Sector. Decis. Anal. 3, (2016).

20. Kozlovskiy, I., Sodenkamp, M., Hopf, K. and Staake, T.: Energy Informatics for

Environmental, Economic and Societal Sustainability: A Case of the Large-Scale Detection of Households with Old Heating Systems. In: Proceedings of the Twenty-Fourth European Conference on Information Sytsems, AISeL, Istanbul, Turkey (2016).

21. Sodenkamp, M., Kozlovskiy, I., Staake, T.: Gaining IS Business Value through Big Data Analytics: A Case Study of the Energy Sector. In: Proceedings of the Thirty Sixth International Conference on Information Systems (ICIS), Fort Worth, USA, 13-16 December 2015. AISeL, Fort Worth, USA (2015).

22. Hopf. K., Sodenkamp, M., Kozlovskiy, I.: Energy Data Analytics for Improved Residential Service Quality and Energy Efficiency. In: Proceedings of the Twenty-Fourth European Conference on Information Sytsems, AISeL, Istanbul, Turkey (2016).

23. U.S. National Centers for Environmental Information: Climate Data Online, http://www.ncdc.noaa.gov/cdo-web/.

24. Sodenkamp M., Hopf, K., Kozlovskiy, I., Staake, T. Smart-Meter-Datenanalyse für automatisierte Energieberatungen ("Smart Grid Data Analytics") - Schlussbericht für das Bundesamt für Energie Schweiz (Final project report), (2016).

25. Romanski, P., Kotthoff, L.: FSelector: Selecting attributes. (2014).

26. Freund, Y., Schapire, R.E.: A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. J. Comput. Syst. Sci. 55, 119–139 (1997).

27. Zaki, M.J., Meira Jr., W.: Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press (2014).

28. Breiman, L.: Random Forests. Mach. Learn. 45, 5–32 (2001).

29. Vapnik, V.N., Vapnik, V.: Statistical Learning Theory. Wiley New York (1998).

30. Cramer, H.: Mathematical Methods of Statistics. Princeton University Press, Princeton (1946).

31. Jurman, G., Riccadonna, S., Furlanello, C.: A Comparison of MCC and CEN Error Measures in Multi-class Prediction. PloS One. 7, (2012).

32. Gorodkin, J.: Comparing two K-Category Assignments by a K-Category Correlation Coefficient. Comput. Biol. Chem. 28, 367–374 (2004).

33. Dietterich, T.G.: Ensemble Methods in Machine Learning. In: International Workshop on Multiple Classifier Systems. pp. 1–15. Springer (2000).

Dead or Alive? A Formal Decision Model for Deciding on Customer Recovery Investments

Dominikus Kleindienst[1] and Daniela Waldmann[1,2]

[1] University of Augsburg, FIM Research Center, Augsburg, Germany
{dominikus.kleindienst,daniela.waldmann}@fim-rc.de
[2] Fraunhofer FIT - Project Group Business and Information Systems Engineering, Augsburg, Germany

Abstract. As digitization makes customer migration easier and more attractive, managing customer recovery becomes increasingly important for organizations. In this context, the challenge is to avoid two error types that can occur with customer relation recovery. First, mistakenly investing in customer relations that are active ("alive"), and, second, mistakenly not investing in migrated customer relations ("dead"). Consequently, considering the probability that a customer relation is "alive" or "dead" is necessary. Based on this probability, an economically reasonable decision has to be made whether to invest in individual customer relation recovery or not. However, existing literature often neglects the above mentioned probability. Accordingly, based on a comprehensive discussion of related work, we propose a formal decision model on whether to invest in customer relation recovery, considering the probability that the customer relation is still "alive" or "dead." To demonstrate the decision model's applicability, an illustrative case with a sample calculation is presented.

Keywords: customer data, customer recovery, digitization, decision model

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Kleindienst, D.; Waldmann, D. (2017): Dead or Alive? A Formal Decision Model for Deciding on Customer Recovery Investments, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1250-1263

1 Introduction

In all types of organizations, customers come and go or, following Schmittlein et al. [1], there are customers who are “alive” (i.e., maintain an active customer relation) and customers who are “dead” (i.e., left the organization for whatever reason). With digitization as a main driver of the global economy, making this distinction is becoming more and more challenging: customer migration becomes easier and more attractive as digitization is breaking down barriers of entry, enables more transparent markets, and is comparatively impersonal [2]. Therefore, the importance of managing customer recovery is increasing. This challenge intensifies in settings where customers are not bound by contracts and have the possibility to change between different vendors [3, 4]. On the other side, digitization does not only amplify customer migration, but can also be an advantage for organizations because of the significant increase in available customer data [5]. Consequently, with the increasing amount of available data, opportunities to identify customer insights from data continue to expand. Organizations have the possibility to collect, store, and analyze available customer data [6] and use it, among others, for custom-designed investment decisions in the recovery of individual customer relations.

The related literature already provides models for deciding between different marketing investment alternatives [7, 8] and develops guidelines for acquisition and retention decisions [9, 10]. However, although economic aspects such as cost-benefit-trade-offs are investigated in the literature, it is not always clear for which customer relation a recovery is actually required. To illustrate this, it would, for example, not be reasonable to invest in recovering a customer relation, which is active and would likely realize further transactions anyway. However, as the point at which a customer relation becomes inactive cannot always be known for sure, it has to be estimated using indicators such as a low number of transactions or an unexpectedly long time since the last transaction [1]. Therefore, if a customer relation is still “alive” can be estimated as a conditional probability, given certain purchase information [1].

In summary, customer recovery faces two challenges. First, organizations need to identify the probability that a customer relation is to be considered “dead” or “alive.” Second, once such probability can be estimated, the question is at which probability an investment in the recovery of an individual customer relation is economically reasonable. Thus, we aim at an analytical solution to the following research question: how can a decision maker decide on investing in a customer relation on the basis of the probability that the customer relation is still “alive“?

To answer this question, we develop a formal decision model based on the following ideas. On one hand, investing in the recovery of a customer relation is only reasonable if the present value of future cash flows when investing is higher than the present value of future cash flows when not investing in the recovery of a customer relation. On the other hand, deciding on investing in customer recovery should consider the probability that a customer relation is “dead” or “alive.” Consequently, we compute a threshold from which on investment in an individual customer relation is economically reasonable.

We structure the paper starting with a discussion of the problem context and related work. Further, we propose the decision model and demonstrate its application by a sample calculation. Lastly, we discuss the resulting decision model.

2 Problem Context

Digitization makes customer focus more valuable for organizations and a bigger challenge, as customer migration becomes easier and more attractive than before [11]. In this environment, organizations do not only need to acquire new customers and build loyalty among existing customers, but also target migrated customers. Typically, customer migration arises as soon as there is a gap between the priorities of the customer and the activities of the organization [12]. From an organization's perspective, the reasons for such a gap can be manifold: the lack of success in identifying and using interesting market opportunities, limited information about competitors, no effective communication with the market, no comprehensive customer service, or missing knowledge about customer needs, their perceptions, preferences, and behavior [13]. In summary, in order to avoid customer migration, organizations need to increase customer satisfaction as it affects customers' repurchase likelihood [14]. Customer satisfaction can, for example, be defined as a cumulative evaluation of a customer's purchase and consumption experience to date [14–17]. Generally, customers are considered "alive" as long as they are still cultivating an active relation to the organization, and "dead" if they terminated their relation to the organization for whatever reason [1]. To recover "dead" customer relations, an organization needs to first identify the respective customers. On the one hand, identifying if a customer relation is "alive" or not can be obvious. For instance, if customers cancel their cell-phone contract and change their provider, the customer relation would clearly be considered "dead." On the other hand, there are contexts in which a customer relation transition from "alive" to "dead" is not always that easy to detect for companies, as customers "may not notify the firm when they leave" [1]. This holds true for hotel stays, air travel, or large online retailers such as Amazon, where customers are not bound by contracts and have the possibility to switch between different vendors [3, 4]. Hence, particularly in non-contractual customer relations, it is a challenge for organizations to know whether a customer relation is "alive" or "dead" [1]. One indicator for companies whether a customer relation is "alive" or "dead" is a customer's purchasing information (e.g. an unexpectedly long time period since the last transaction). However, a long transaction break does not necessarily mean that a customer relation is definitively "dead" [18].

Generally, related literature approaches this topic by modeling customer migration. Several researchers use recency in models that predict customer behavior. For example, Bult and Wansbeek [19], Bitran and Mondschein [20], Fader et al. [18], and Rhee and McIntyre [21] find a negative association between recency and purchase likelihood. Dwyer [22] identifies "always-a-share" customers' purchase probability by developing a purchase decision-making tree based on historical buying data. The Dwyer model is used in most customer lifetime value (CLV) research [23]. Comprehensive explanations on the CLV can be found in Kumar and Reinartz [24] for instance. Blattberg et al. [10] extend the Dwyer model and use the “recency, frequency, monetary index” to develop the purchase decision-making tree [23]. In brief, literature has long contributed to the understanding of customer migration and the factors affecting it. However, at this point, no concrete implications for customer recovery investments have been derived.

Customer recovery campaigns are a specific kind of customer campaigns. According to the campaign management process of Englbrecht [25], campaigns are mainly characterized by target group, channel, and content. Hence, investment decisions are to be made between different campaign alternatives comprising possible target groups, channels, or content. Here, the target group comprises migrated customers, or in other words “dead” customer relations. Channels, for instance, are categorized into offline channels, such as stores or catalogues, and online channels, such as mobile apps, email, or websites. They can also be differentiated by direct and indirect channels, distinguished by whether there is an intermediary responsible for managing the relationship between the customer and the organization [26]. Typically, the content of customer recovery campaigns can entail special offers, discounts, vouchers, coupons or other incentives for recovering customer relations.

For deciding on competing marketing investments, the literature provides numerous approaches. For example, Rust et al. [7] provide a framework to trade off competing marketing investments on the basis of financial return. Neslin et al. [8] demonstrate how to target the right marketing to the right customers at the right time to maximize CLV. Venkatesan and Kumar [27] recommend CLV as a metric for selecting customers and designing marketing programs, as they provide empirical evidence for the existence of a relationship between marketing actions and CLV. As such, Glady et al. [28] show that the dependence between the number of transactions and their profitability can be used to increase the accuracy of the CLV. Venkatesan and Kumar [27] point out that literature provides guidelines for acquisition and retention decisions [9, 10]. There are also studies on the basis of which customers should be “eliminated”: Reinartz and Kumar [29], for instance, demonstrate how to decide on terminating a customer relation or not. In summary, literature provides various discussions and models concerning investments in customer relations. However, to the best of our knowledge, there is no formal decision model on the economic feasibility of customer recovery investments while considering the probability that a customer relation is actually “dead.”

In order to fill this gap, we introduce a calculation to identify the most economically reasonable investment alternative out of multiple customer recovery investments, and propose a decision model that advises whether to invest in customer recovery or not. Based on an existing decision model designed to manage data currency [30], we premise our model on the probability that a customer relation is still “alive.“ The detailed decision model is described in the subsequent chapter.

3 Decision Model

The basic idea of the model is to decide on investing in customer recovery or not by comparing a threshold from which an investment is economically useful to the current probability that a customer relation is "alive." This threshold depends on investment specific variables, such as the costs of the investment and an effectiveness factor, and customer specific ones, such as the present value of future cash inflows of a customer relation. Beside, we make some assumptions in our decision model. First, the decision model is designed to cover a single period. Second, we assume that the organization's risk attitude is neutral when deciding on customer recovery investments. Third, for reasons of simplicity, we only distinguish between "dead" and "alive" customers and do not consider possible attributes in between.

The decision model has four steps: (1) selection of the most economically reasonable investment alternative; (2) determination of the current probability that a customer relation is "alive;" (3) derivation of the threshold; and (4) making the investment decision.

Step 1: Selection of the most economically reasonable investment alternative

Organizations have to decide between several investment alternatives for customer recovery. As such, the decision underlies the expected cash flow $E(CF_{ij}) \in \mathbb{R}_0^+$ of a customer relation i $(i = 1, \dots, n \; with \; n \in \mathbb{N})$ when successfully recovering it with a specific investment alternative j $(j = 1, \dots, m \; with \; m \in \mathbb{N})$. To calculate $E(CF_{ij})$, the present value of future cash inflows of a customer relation $\pi_i \in \mathbb{R}^+$, the investment costs $I_j \in \mathbb{R}^+$, and the effectiveness factor $\eta_{ij} \in (0; 1]$, which determine the success probability of a recovery investment, are necessary as they all influence the economic assessment of the investment alternatives. The domain of π_i is defined as $\mathbb{R}^+$, as only customers with positive cash flows are of interest. The domain of η_{ij} excludes the value 0, as we exclude investment alternatives for which customer recovery is impossible. Additionally, investments in customer relations with negative expected cash flows $E(CF_{ij})$ are not economically reasonable. Therefore, $E(CF_{ij})$ is only defined for $\pi_i \cdot \eta_{ij} - I_j \geq 0$. Hence, the calculations represented by equation 1 lead to the economically optimal investment alternatives J^*.

$$J_i^* = \{j \in (1, \dots, m); \forall k \in (1, \dots, m) \setminus \{j\} : \; E(CF_{ik}) \leq E(CF_{ij})\},$$
$$\text{with } E(CF_{ij}) = \pi_i \cdot \eta_{ij} - I_j, \tag{1}$$

where J_i^* represents the set of all indices j for which the expected cash flow $E(CF_{ij})$ of a customer relation i for a specific investment alternative j is maximal. In case of multiple resulting indices j, that is, indices j with the same expected cash flows, $E(CF_{ij})$, the decider should take the investment alternative j that is cheaper after normalization to effectiveness (e.g., if $J_i^* = \{1,2\}$ and $I_1 < I_2 \cdot \frac{\eta_{i1}}{\eta_{i2}}$, then decide for $j = 1$).

Step 2: Measuring the probability that a customer relation is "alive"

In the following, we use the model for assessing conditional probability of Schmittlein et al. [1] as a basis for estimating the probability that a customer relation i is "alive." According to Fader et al. [31], the model of Schmittlein et al. [1] shows an impressive predictive performance, its empirical validation is often presented, and there are several applications in different contexts, such as customer profitability, churn prediction, and customer base analysis [4, 18, 29, 32–35]. The conditional probability $P_i("alive"|Information) \in [0;1]$ depends on a customer's individual purchasing information [1]. This can be $Information = x, t_x, T$, where x is the number of transactions observed in the time interval $(0, T]$ and $t_x (0 < t_x \leq T)$ is the time of the last transaction [1]. That means that recency and frequency are sufficient statistics for an individual customer's purchasing behavior [18]. $P_i("alive"|Information)$ represents the actual probability that a customer relation i with an observed behavior is still "alive," which should be compared to the threshold in order to make an economically reasonable investment decision.

Step 3: Threshold derivation

As it is highly improbable to know for sure if a customer relation is "alive" or "dead," there is always the possibility that the organization comes to a "correct" or "wrong" investment decision for a customer relation i. Regarding "wrong" investment decisions, it is possible that organizations either unnecessarily invest in "alive" customer relations (see Table 1, case Ia) or do not invest in "dead" customer relations with positive expected cash flows in case of customer relation recovery (see Table 1, case Ib). By taking such "wrong" decisions, the organization either unnecessarily loses investment costs or cash inflows which might result from investment j. Accordingly, cases IIa and IIb represent "correct" decisions as long as $\pi_i \cdot \eta_{ij} \geq I_j$. Table 1 represents all possibilities of total expected cash flows depending on $P_i("alive"|Information)$ and the decisions to invest in customer relations or not.

Table 1. Matrix of the total expected cash flow

Decision	*"Dead"* $1 - P_i("alive"\|Information)$		*"Alive"* $P_i("alive"\|Information)$	
Investment	IIa	$\pi_i \cdot \eta_{ij} - I_j$	Ia	$\pi_i - I_j$
No investment	Ib	$-\pi_i \cdot \eta_{ij}$	IIb	π_i

The total expected cash flow of case Ia represents the present value of future cash inflows π_i resulting from a customer relation i minus the investment costs I_j of investment alternative j for investing in a customer relation (see equation 2). Here, the investment costs I_j arise unnecessarily.

$$E(CF_{ij})(investment \wedge "alive") = \pi_i - I_j. \tag{2}$$

In contrast, case Ib leads to a lost present value of future cash inflows $-\pi_i$ caused by not recovering a customer relation i. This lost present value of future cash inflows, which corresponds to opportunity costs, only comes into force to the extent to which the customer recovery investment would have been successful, which is represented by the effectiveness factor η_{ij}:

$$E(CF_{ij})(no\ investment \wedge "dead") = -\pi_i \cdot \eta_{ij}. \tag{3}$$

Given $\pi_i \cdot \eta_j \geq I_j$, investing in a “dead” customer relation and not investing in an “alive” one are correct decisions. Hence, case IIa entails the present value of future cash inflows of a customer relation π_i multiplied with the effectiveness factor η_{ij}, reduced by the costs of investment I_j:

$$E(CF_{ij})(investment \wedge "dead") = \pi_{\boldsymbol{i}} \cdot \eta_{ij} - I_j. \tag{4}$$

Case IIb represents not investing in an “alive” customer relation, which results in the present value of future cash inflows $\pi_{\boldsymbol{i}}$:

$$E(CF_{ij})(no\ investment \wedge "alive") = \pi_i. \tag{5}$$

Based on these mathematical terms, the threshold for an economic decision on whether to invest in a customer relation or not can be deduced. From an economic point of view, investing in a customer relation is only reasonable if the total expected cash flow in case of an investment for recovering a customer relation is higher than the total expected cash flow for not investing (see equation 6). The cases Ia, Ib, IIa, and IIb arise with the probabilities that a customer relation i is “alive,”$P_i("alive"|Information)$, or already “dead,” $1 - P_i("alive"|Information)$, as presented in Table 1. Equation 6 covers decisions under risk neutral preferences:

$$\begin{gathered} E(CF_{ij})(investment \wedge "alive") \cdot P_i("alive"|Information)\ + \\ E(CF_{ij})(investment \wedge "dead") \cdot (1 - P_i("alive"|Information))\ > \\ E(CF_{ij})(no\ investment \wedge "alive") \cdot P_i("alive"|Information) + \\ E(CF_{ij})(no\ investment \wedge "dead") \cdot (1 - P_i("alive"|Information)). \end{gathered} \tag{6}$$

After computing terms 2–5, we solve the inequality for $P_i("alive"|Information)$ (see equation 7), which results in the threshold $T_{ij} \in [0;1)$:

$$T_{ij} > P_i(\text{"}alive\text{"}|Information),$$

$$with\ T_{ij} = \frac{2\pi_i\eta_{ij} - I_j}{2\pi_i\eta_{ij}}. \tag{7}$$

The threshold enables making investment decisions in which the total expected cash flow in case of an investment is higher than the total expected cash flow for not investing in recovering a customer relation.

Step 4: Making the investment decision

To make the investment decision D_i for a customer relation i, the organization should now compare the probability that the customer relation is still "alive," $P_i(\text{"}alive\text{"}|Information)$, with the threshold T_{ij}:

$$D_i = \begin{cases} invest & for\ T_{ij} > P_i(\text{"}alive\text{"}|Information) \\ not\ invest & for\ T_{ij} \leq P_i(\text{"}alive\text{"}|Information) \end{cases}. \tag{8}$$

In summary, the four proposed steps lead to an economically reasonable decision on whether to invest in an individual customer relation's recovery or not by comparing the threshold to the current probability that a customer relation is "alive," as per equation 8.

4 Application and evaluation

We illustrate the applicability, completeness, understandability, feasibility, and operationality of the decision model by an example in which an online retailer aims at recovering possibly "dead" customer relations. At the same time, the online retailer wants to avoid unnecessarily investing in "alive" customer relations. By using our decision model, the online retailer addresses only those customer relations for which an investment is reasonable on the basis of the probability that they are "alive" compared to the calculated threshold. As such, we show the economic benefit of the decision model.

Step 1: Selection of the most economically reasonable investment alternative

At first, the online retailer has to identity different investment alternatives for customer recovery and select the most economically reasonable investment alternative for every customer relation i. In our example, the online retailer selects four possible investment alternatives j, that is, two different channels, letter and mail, and two different contents, voucher and special offer. According to the experience of the online retailer, customer recovery via letter is more effective than email, and a voucher is more effective than a special offer. Moreover, in this example, customer recovery with vouchers incurs more investment costs than special offers. Table 2 shows the effectiveness factor η_{ij} and the costs of the four investment alternatives I_j.

Table 2. η_{ij} and I_j for the investment alternatives

		Special offer via letter ($j = 1$)	*Voucher via letter* ($j = 2$)	*Special offer via email* ($j = 3$)	*Voucher via email* ($j = 4$)
η_{ij}	$i = 1$	0,09	0,05	0,17	0,08
	$i = 2$	0,08	0,06	0,08	0,17
	$i = 3$	0,06	0,19	0,08	0,15
I_j		USD 20	USD 30	USD 12	USD 22

To select the most economically reasonable investment alternatives for different customers according to Formula 1, we take the present values of future cash inflows π_i of three customers as example: $\pi_1 = USD\ 640$, $\pi_2 = USD\ 857$, and $\pi_3 = USD\ 428$. Table 3 shows the expected cash flows for each customer relation i and the four different investment alternatives j.

Table 3. $E(CF_{ij})$ for the customer relations and investment (USD)

		Investment alternatives			
		$j = 1$	$j = 2$	$j = 3$	$j = 4$
Customer Relations	$i = 1$	35.89	3.52	**97.07**	32.33
	$i = 2$	45.75	20.96	55.46	**126.67**
	$i = 3$	7.33	**52.06**	21.98	40.09

The results of Table 3 show that the most economically reasonable investment alternative for the customer relation $i = 1$ is $j = 3$, for $i = 2$ is $j = 4$, and for $i = 3$ is $j = 2$ (see bold marked values in Table 3), as these investment alternatives have the greatest expected cash flow for the different customers as per equation 1.

Step 2: Measuring the probability that a customer relation is "alive"

Next, the online retailer has to quantify the probability that a customer relation is still "alive," $P_i(\text{"}alive\text{"}|Information)$. For instance, the organization can follow Schmittlein and Peterson [32] and Reinartz and Kumar [4], who determine the probability depending on recency and transaction frequency. For example, we assume the following values for the three customers: $P_1(\text{"}alive\text{"}|Information) = 0.30$, $P_2(\text{"}alive\text{"}|Information) = 0.96$, and $P_3(\text{"}alive\text{"}|Information) = 0.23$.

Step 3: Calculation of the threshold

Further, the online retailer has to calculate the threshold for the customers and the selected investment alternative by using equation 7.

Table 4. Results of the threshold T_{ij}

	$J_1^* = 3$	$J_2^* = 4$	$J_3^* = 2$
T_{ij}	94.40%	92.60%	81.72%

Step 4: Making the investment decision

The application of equation 8 shows whether the online retailer should invest in the customer relations or not by comparing the threshold T_{ij} with the probability that the customer relation i is still "alive," $P_i("alive"|Information)$.

Table 5. Investment decision for the customer relations i

			D_i
$i = 1$	T_{13}	94.40%	invest
	$P_1("alive"\|Information)$	30.00%	
$i = 2$	T_{24}	92.60%	not invest
	$P_2("alive"\|Information)$	96.00%	
$i = 3$	T_{32}	81.72%	invest
	$P_3("alive"\|Information)$	23.00%	

Table 5 shows that the online retailer should invest in customer relation $i = 1$ and $i = 3$ because the results of the threshold T_{13} and T_{32} are greater than $P_1("alive"|Information)$ and $P_3("alive"|Information)$, respectively. For customer relation $i = 2$ the investment decision is not to invest, as T_{24} is less than $P_2("alive"|Information)$.

Next, we explain why the evaluation criteria is fulfilled. First, applicability is shown by conducting this example calculation. Further, this evaluation type demonstrates the decision model's completeness as all input variables are comprehensive and quantitative measures. The evaluation criterion understandability is shown as the actual measure is easy to interpret and applicable for users. Feasibility and operationality is given as the parameters are determinable, well defined, and the decision model is based on a quantitative measurement. Additionally, the data necessary for the model calculation are accessible and affordable, as the number of transactions or the time of the last transaction usually exist in organizations. In order to evaluate the decision model from an economic perspective, we extend the sample calculation and instantiate the decision model with 10,000 customer relations that have equally distributed probabilities of being "alive" in an interval of 0–100% and equally distributed expected cash flows in an interval of USD 0–1,000. The effectiveness factors of the investment alternatives are equally distributed in an interval of 0–20%. In this sample calculation, we use the parameter setting of the four investment alternatives listed in Table 2. In case of a perfect estimation of the probabilities that customer relations are still “alive,” $P_i("alive"|Information)$, the sample calculation reveals that about 25% of individual recovery investments can be saved, which leads to significant cost savings by applying the decision model. As shown in Table 6, a sensitivity analysis suggests that – on the one hand – estimation errors of expected cash flows and effectiveness factors only lead to disproportionately low changes in cost savings. That is, the model can said to be robust in terms of these parameters.

Table 6. *Impacts of estimation errors on cost savings*

Parameter estimation error	***-20%***	***-15%***	***-10%***	***-5%***	***0%***	***+5%***	***+10%***	***+15%***	***+20%***
$\boldsymbol{\eta_{ij}}$	+1%	0%	0%	0%	0%	-1%	-1%	-2%	-2%
$\boldsymbol{E(CF_{ij})}$	+6%	+5%	+3%	+1%	0%	-2%	-4%	-6%	-7%
$\boldsymbol{P_i("alive"|Inf.)}$	-82%	-71%	-56%	-29%	0%	+25%	+48%	+72%	+92%

On the other hand, the sensitivity analysis exposes that $P_i("alive"|Information)$ has to be estimated carefully, as estimation errors lead to disproportionately high changes in cost savings. However, even for poorer estimations, savings on a low percentage basis can be generated, that easily be significant in monetary terms for large customer recovery investments.

5 Summary and Discussion

In this paper, we point out that, for customer recovery, distinguishing between “alive” and “dead” customer relations becomes a challenge with the increase in market transparency and impersonality. Therefore, organizations risk wrong investment decisions when managing customer recovery. Addressing this challenge, literature offers solutions to determine the economic value of a customer relation as a basis for this decision. However, to the best of our knowledge, no approach considers the probability that a customer relation is “alive” for such investment decisions. Therefore, we combine these ideas in a formal decision model for deciding on customer recovery in an economically reasonable manner by considering the probability that a customer relation is “dead.” In doing so, we strive for practical applicability and demonstrate the decision model’s operationalization in an illustrative example.

Nevertheless, our decision model has limitations that stimulate further research. First, further research should examine the decision model in a real world context in order to evaluate its usefulness [36]. However, we can evaluate the decision model in terms of its applicability, completeness, understandability, feasibility, and operationality by using an example. In doing so, we follow Sonnenberg and Vom Brocke [36], and argue that it is reasonable to disseminate research findings in early stages to communicate them to interested peers and research communities. Second, the decision model is designed to cover a single period. In practice, in order to permanently ensure maximum of customer recovery, periodical assessments could be a possible extension of the decision model. Third, in the decision model it is assumed that the organization's decision regarding customer recovery investments is risk neutral. In reality, risk attitude can be context and branch specific, which should be further examined in future research. Fourth, future research can unfold the range between the attributes "dead" and "alive" to further approach the decision model for context- and industry-specific dependencies due to individual stages in the customer lifecycle. Finally, we assume that investing in an "alive" customer relation is not reasonable in terms of a recovery effect. In reality, recovery investments could also increase the satisfaction of "alive" customer relations.

However, besides these limitations and the identified need for further investigation, we consider the presented approach to be a valuable contribution to research and practice in order to enable data-driven recovery investment decisions. Accordingly, whether to invest in recovering a customer relation can be decided on the basis of a substantiated formal decision model. Moreover, by using digitally available customer data in the proposed decision model, organizations have the possibility to better meet digitization challenges in customer recovery.

References

1. Schmittlein, D.C., Morrison, D.G., Colombo, R.: Counting your customers. Who are they and what will they do next? Management Science 33, 1–24 (1987)
2. Desai, D.R.: The new steam. On digitization, decentralization, and disruption. Hastings Law Journal 65, 1469–1482 (2014)
3. Dwyer, R.F.: Customer lifetime valuation to support marketing decision making. Journal of Direct Marketing 3, 8–15 (1989)
4. Reinartz, W.J., Kumar, V.: On the profitability of long-life customers in a noncontractual setting. An empirical investigation and implications for marketing. Journal of Marketing 64, 17–35 (2000)
5. Mayer-Schönberger, V., Cukier, K.: Big data. A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt, New York, New York (2013)
6. Beath, C., Becerra-Fernandez, I., Ross, J., Short, J.: Finding value in the information explosion. MIT Sloan Management Review 53, 18–20 (2012)
7. Rust, R.T., Lemon, K.N., Zeithaml, V.A.: Return on marketing. Using customer equity to focus marketing strategy. Journal of Marketing 68, 109–127 (2004)
8. Neslin, S.A., Taylor, G.A., Grantham, K.D., McNeil, K.R.: Overcoming the "recency trap" in customer relationship management. Journal of the Academy of Marketing Science 41, 320–337 (2013)
9. Blattberg, R.C., Deighton, J.: Manage marketing by the customer equity test. Journal of Interactive Marketing 74, 136–144 (1996)
10. Blattberg, R.C., Getz, G., Thomas, J.S.: Customer equity. Building and managing relationships as valuable assets. Harvard Business School Press, Boston, Mass. (2001)

11. Rezabakhsh, B., Bornemann, D., Hansen, U., Schrader, U.: Consumer power: A comparison of the old economy and the internet economy. Journal of Consumer Policy 29, 3–16 (2006)
12. Kotler, P.: Ten deadly marketing sins. Signs and Solutions. Wiley, Hoboken, New York (2004)
13. Czarniewski, S.: Changes in consumer behaviour in the market and the value of companies. European Journal of Research and Reflection in Management Sciences 2, 61–68 (2014)
14. Auh, S., Johnson, M.D.: Compatibility effects in evaluations of satisfaction and loyalty. Journal of Economic Psychology 26, 35–57 (2005)
15. Johnson, M.D., Anderson, E.W., Fornell, C.: Rational and adaptive performance expectations in a customer satisfaction framework. Journal of Customer Research 21, 128–140 (1995)
16. Lervik, O.L., Johnson, M.D.: Service equity, satisfaction and loyalty. From transaction-specific to cumulative evaluations. Journal of Service Research 5, 184–195 (2003)
17. Rust, R.T., Zahorik, A.J., Keiningham, T.L.: Return on quality (ROQ): Making service quality financially accountable. The Journal of Marketing 59, 58–70 (1995)
18. Fader, P.S., Hardie, B.G.S., Lok Lee, K.: RFM and CLV. Using iso-value curves for customer base analysis. Journal of Marketing Research 42, 415–430 (2005)
19. Bult, J.R., Wansbeek, T.: Optimal selection for direct mail. Marketing Science 14, 378–394 (1995)
20. Bitran, G.R., Mondschein, S.V.: Mailing decisions in the catalog sales industry. Management Science 42, 1364–1381 (1996)
21. Rhee, S., McIntyre, S.: Including the effects of prior and recent contact effort in a customer scoring model for database marketing. Journal of the Academy of Marketing Science 36, 538–551 (2008)
22. Dwyer, R.F.: Customer lifetime valuation to support marketing decision making. Journal of Direct Marketing 11, 6–13 (1997)
23. Qi, J., Shu, H., Li, H.: Study on purchase probability model in CRM systems. Research and Practical Issues of Enterprise Information Systems, 643–647 (2006)
24. Kumar, V., Reinartz, W.J.: Customer relationship management. Concept, strategy, and tools. Springer, Berlin (2012)
25. Englbrecht, A.: Kundenwertorientiertes Kampagnenmanagement im CRM. Kovač, Hamburg (2007)
26. Hosseini, S., Oberländer, A., Röglinger, M., Wolf, T.: Rethinking multichannel management in a digital world – A decision model for service providers. In: Proceedings of the 12th International Conference on Wirtschaftsinformatik (WI), Osnabrück, Germany, March 2015 (2015)
27. Venkatesan, R., Kumar, V.: A customer lifetime value framework for customer selection and resource allocation strategy. Journal of Marketing 68, 106–125 (2004)
28. Glady, N., Baesens, B., Croux, C.: A modified Pareto/NBD approach for predicting customer lifetime value. Expert Systems with Applications 36, 2062–2071 (2009)

29. Reinartz, W.J., Kumar, V.: The impact of customer relationship characteristics on profitable lifetime duration. Journal of Marketing 67, 79–99 (2003)
30. Görz, Q.: An economics-driven decision model for data quality improvement – A contribution to data currency. In: Proceedings of the 17th Americas Conference on Information Systems, Detroit, Michigan, August 2011 (2011)
31. Fader, P.S., Hardie, B.G.S.: Probability models for customer-base analysis. Journal of Interactive Marketing 23, 61–69 (2009)
32. Schmittlein, D.C., Peterson, R.A.: Customer base analysis. An industrial purchase process application. Marketing Science 13, 41–67 (1994)
33. Hopmann, J., Thede, A.: Applicability of customer churn forecasts in a non-contractual setting. In: Innovations in Classification, Data Science, and Information Systems. Springer, Berlin (2005)
34. Wübben, M., Wangenheim, F.: Instant customer base analysis: Managerial heuristics often "get it right". Journal of Marketing 72, 82–93 (2008)
35. Zitzlsperger, D.F., Robbert, T., Roth, S.: Forecasting customer buying behaviour: The impact of "one-time buyer". In: Proceedings of the ANZMAC 2007 Conference, University of Otago, Dunedin, New Zealand, December 2007 (2007)
36. Sonnenberg, C., Vom Brocke, J.: Evaluations in the science of the artificial – Reconsidering the build-evaluate pattern in design science research. In: Peffers, K., Rothenberger, M., Kuechler, B. (eds.) Design Science Research in Information Systems, pp. 381–397. Springer, Berlin (2012)

Understanding Consumer Behavior in Electronic Commerce with Image Sentiment

Christof Naumzik[1], Stefan Feuerriegel[1], and Dirk Neumann[2]

[1] University of Freiburg, Freiburg, Germany
{christof.naumzik,stefan.feuerriegel,dirk.neumann}@is.uni-freiburg.de

1 Introduction

Customers in electronic commerce translate available information into buying decisions based on needs and behavioural tendencies. For this purpose, shopping sites provide detailed information on products and services ranging from simple figures (e.g., dimensions) to textual descriptions, user-generated reviews and visual images. All of the aforementioned sources of information may have a substantial impact on the decision-making of customers.However, little is known about the ways in which visual information is processed and used in decision-making. As a remedy, this research seeks to shed light on the informativeness of visual content in electronic commerce. This objective represents a highly relevant area of research, since visual information plays a major role in determining price, choice and thus willingness to buy (see [1], [2]). Unfortunately, such an undertaking is challenging as we are not aware of any existing methods for measuring the degree of informativeness of images. The main focus of visual sentiment analysis has been on the classification of the sentiment of images stemming from a fairly heterogeneous set. In contrast, we aim to build a model addressing a somewhat homogeneous set of images. In our case, we use photographs from real estate listings in order to predict the rent price. As an immediate implication, our findings show operators of e-commerce or recommender systems how to optimize the presentation of their products and services.

2 Dataset

In our paper, we demonstrate the added value of visual cues in electronic commerce. More precisely, we investigate how images in apartment listings can work as a predictor of the corresponding rent price. For that purpose, we utilize the data as follows: our dataset consists of 2,500 apartment listings for the Boston area, which we collected from Craigslist1 in June 2016. For each listing, we retrieved both the monthly rent and the size of the apartment, as well as the first picture shown on the webpage.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Naumzik, C.; Feuerriegel, S.; Neumann, D. (2017): Understanding Consumer Behavior in Electronic Commerce with Image Sentiment, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1264-1266

3 Method Development

Figure 1 illustrates our research framework. We train different machine learning models to predict rent prices based merely on the image content. We chose to utilize both a Support Vector Regression (SVR) and a Random Forest (RF), as these models are already suitable for smaller datasets and allow for comparatively fast computation. To evaluate the predictive performance of our proposed method, we randomly partition our dataset into two subsets, a training set of 1800 and a test set of the remaining 700 listings.

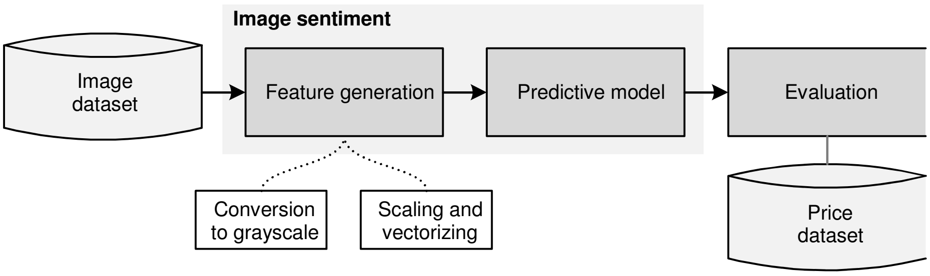


Figure 1. Research framework shows the steps of preprocessing for feature generation and the subsequent predictive model in order to calculate a sentiment metric for images.

4 Empirical Findings

The research framework shows the steps of preprocessing for feature generation and the subsequent predictive model used to calculate a sentiment metric for images. In table 1, we present the results from our predictive approach in comparison to our benchmark. We note that the SVR, as well as the Random Forest model, outperforms the benchmark by a substantial margin for both RMSE and MAE. The best results are obtained in the case of the random forest model, with improved prediction accuracy in terms of RMSE and MAE of 50.81% and 62.96 %, respectively. To provide a statistical underpinning for our results, we also performed a Diebold-Mariano test.

5 Contribution and Implications

This research paper contributes to our understanding of human behaviour in electronic commerce as well as, for instance, recommender systems. We propose an innovative approach to scoring the polarity of an image in a computerized fashion and introduce the concept of image sentiment, which refers to the positive or negative characteristics of the object in the image. In contrast to previous research, we avoid judgment as to whether humans perceive the object itself in a positive or negative manner. Instead, we suggest a sentiment measure that ranks images depicting the same object on a continuous scale from a low to a high appraisal.

We have implemented a predictive approach to forecast the price of objects on the basis of their product image. The above case study demonstrates that our predictions can considerably outperform a naïve baseline. This approach can be a viable extension to existing models that describe real estate prices, since the image sentiment is likely to make up for a substantial portion of the error term.

Overall, our research suggests numerous new applications, each of which can yield an immediate impact for professionals. Operators of e-commerce or recommender systems can leverage our method to optimize the presentation of their products and services. Such advances reveal the potential benefits and insights that can be derived from unstructured information.

Table 1. A validation set is used to assess the performance of our trained models. We compare these in terms of (1) RMSE and (2) MAE with the price per square foot in our validation set. Furthermore, we apply a Diebold-Mariano (DM) test to the trained models and the sample mean to test for the same forecasting accuracy.

Method	**RMSE**	**MAE**	**DM test** (test statistic)	**DM test** (p-value)
Benchmark				
Sample mean	1.1159	0.8549	—	—
Predictive models				
SVR (linear kernel)	0.9536	0.6006	-2.3048	0.0216*
	-14.54%	-29.74%		
SVR (radial kernel)	0.8907	0.5572	-4.3236	0.0001***
	-20.18%	-34.82%		
Random forest	0.5489	0.3166	-7.137	0.0001***
	-50.81%	-62.96%		

Statistical significance levels: *** 0.001, ** 0.01, * 0.05

References

1. Cyr, D., Head, M., Larios, H., Pan, B.: Exploring Human Images in Website Design: A Multi-Method Approach. MIS Quarterly 33, 539–566 (2009)
2. Gaykema, F., Burns, P.D. (eds.): Assessing product image quality for online shopping. SPIE (2012)

Over-Paid Search: When Bricks-and-Mortar Retailers Should Not Use Paid Search

Darius Schlangenotto[1], Dennis Kundisch[1]

[1] Paderborn University, Faculty of Business Administration and Economics, Paderborn, Germany
{Darius.Schlangenotto,Dennis.Kundisch}@uni-paderborn.de

Abstract. Current research on paid search highlights its ability to enhance online as well as offline sales but is limited to pure online players and multi-channel firms. This paper presents a controlled field experiment which investigates whether paid search can increase sales for bricks-and-mortar retailers who solely sell their products via local stores and rely on informational websites to reach their customers. Results suggest that paid search increases the number of potential customers by enhancing the reach of online marketing initiatives. Yet, using a difference-in-differences analysis, our findings show that paid search fails to increase offline sales. Local store customers primarily use paid search as a navigational shortcut to the chain's website which they would have reached even without being exposed to paid search. Consequently, bricks-and-mortar retailers should approach paid search cautiously: whilst it can enhance the reach of marketing initiatives, it seems to have little effect on improving offline purchases.

Keywords: paid search, cross-channel behavior, substitution effects, field experiment, bricks-and-mortar

1 Motivation and Research Question

In this paper we investigate the question of whether paid searches are able to increase sales in local stores. Today, paid search – the mechanism of placing online ads in response to user search queries on search engine result pages (SERP) – is already the main source of Internet advertising revenue and expected to grow by 10% annually over the next four years [1].

As an emerging technology, paid search has spawned numerous new avenues for research especially in the fields of Information Systems and Marketing [2]. The current literature comprises, on the one hand, analytical studies, which focus on the paid search market as a whole and, on the other, empirical studies, which address the benefits of paid search for advertisers [3]. In the empirical stream of literature scholars either studied long-term effects of paid search such as brand perception [4] or short-term effects such as click-behavior of paid search users [5]. Investigating short-term effects, several studies focus on whether paid search has been able to increase online sales, so-called *own-channel effects*. The current body of literature suggests that paid search can be seen as a well suited mechanism to increase the reach of marketing campaigns [6]

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Schlangenotto, D.; Kundisch, D. (2017): Over-Paid Search: When Bricks-and-Mortar Retailers Should Not Use Paid Search, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1267-1270

and enhance online sales [7]. Up to now only a few studies [8] investigated the impact of paid search on offline sales, or so-called *cross-channel effects*. Early results suggest that paid search has the potential of enhancing offline purchases due to consumers' tendency to undertake online research prior to shopping in stores, so-called *research-shopping* [9]. However, current research on the cross-channel effects of paid search has so far only investigated the case of multi-channel firms that offer both online and offline shopping. In contrast, we are interested in finding out whether paid search initiatives could also benefit bricks-and-mortar chains that solely sell in-store and rely on their websites to promote their offerings. If paid search could be shown to have a positive effect on local store purchases, then bricks-and-mortar firms ought to consider investing in paid search in order to increase the reach of their marketing campaigns. With our study we intend to contribute to the current body of literature on paid search by analyzing the benefits of paid search for bricks-and-mortar firms. Hence, we pose the research question: *Do paid search campaigns pay off for bricks-and-mortar retailers?*

2 Methodology

To answer our research question we teamed up with a well-known bricks-and-mortar business-to-consumer (B2C) furniture retailer operating in Germany. The chain's website is primarily used to advertise the products sold in their six local stores. To quantify the cross-channel effect induced by paid search, one needs to address the endogeneity of marketing campaigns [10] as well as limited tracking possibilities between the online and the offline environment. To cope with both problems, in 2015 we conducted a field experiment applying a matched subject design while isolating the cross-channel effect of paid search using a trackable coupon. To identify the causal impact of paid search, we crafted an advertising campaign with Google AdWords in combination with geo-targeting [11] to ensure an appropriate experimental design setting. Throughout the experiment customers in the treatment regions were exposed to paid ads as opposed to customers belonging to the control regions who were not exposed to any paid ads. Based on 85,649 website visits we performed difference-in-differences analyses [12] to estimate paid search impact on user behavior in the online as well as the offline environment.

3 Results

Experimental results suggest that consumers tend to perform research-shopping by using the Internet as a search channel and the local store as a purchase channel. Yet, contrary to current literature [8], [13], our experiment reveals that at least in our case paid search fails to increase the number of cross-channel shoppers. In line with Blake et al. [6] who performed a similar experiment in a pure online environment, we argue that for a well-known bricks-and-mortar chain paid search is partly used by consumers as a navigational shortcut to the advertisers' website, and that cross-channel shopping is primarily performed by these customers. In our study, a paid search campaign is shown to increase the reach of online marketing initiatives, exposing additional customers to the offers, but these additional customers do not engage in cross-channel shopping. In other words, cross-channel shopping is primarily performed by customers who would have been exposed to the offers anyhow, without the firm's investment in paid search. With paid search shown as failing to increase sales, we are able to provide first evidence that it might has no direct sales impact for pure offline retailers.

For managers of bricks-and-mortar firms our study offers various practical implications. First, consumers do perform research-shopping and show cross-channel behavior by using an online coupon for offline purchases. This suggests that online offers might indeed increase local store sales. Second, paid search is able to broaden the reach of marketing campaigns as additional potential customers become aware of the chain's offerings. However, our results suggest that managers who mainly want to increase offline sales should approach an investment in paid search carefully. At least in our case paid search only serves as a costly substitute used by customers who would have performed cross-channel shopping even without being exposed to those ads.

References

1. eMarketer: Advertising Forecast, http://www.emarketer.com/Article/Google-Will-Take-55-of-Search-Ad-Dollars-Globally-2015/1012294 (Accessed: 01.12.2016)
2. Rutz, O., Bucklin, R.: Paid Search Advertising. In: Bock, K.W. de, Neslin, S.A., Coussement, K. (eds.) Advanced Database Marketing. Innovative Methodologies and Applications for Managing Customer Relationships, pp. 229–245. Ashgate Publishing Ltd, Farnham (2013)
3. Desai, P.S., Shin, W., Staelin, R.: The Company That You Keep. When to Buy a Competitor's Keyword. Marketing Science 33, 485–508 (2014)
4. Zenetti, G., Bijmolt, Tammo H. A., Leeflang, Peter S. H., Klapper, D.: Search Engine Advertising Effectiveness in a Multimedia Campaign. International Journal of Electronic Commerce 18, 7–38 (2014)
5. Chan, D.X., Yuan, Y., Koehler, J., Kumar, D.: Incremental Clicks. The Impact of Search Advertising. Journal of Advertising Research 51, 643–647 (2011)
6. Blake, T., Nosko, C., Tadelis, S.: Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment. Econometrica 83, 155–174 (2015)
7. Lu, X., Zhao, X.: Differential Effects of Keyword Selection in Search Engine Advertising on Direct and Indirect Sales. Journal of Management Information Systems 30, 299–326 (2014)
8. Dinner, I.M., Van Heerde, Harald J., Neslin, S.A.: Driving Online and Offline Sales: The Cross-Channel Effects of Traditional, Online Display, and Paid Search Advertising. Journal of Marketing Research 51, 527–545 (2014)
9. Verhoef, P.C., Neslin, S.A., Vroomen, B.: Multichannel customer management. Understanding the research-shopper phenomenon. International Journal of Research in Marketing. 24, 129–148 (2007)
10. Rossi, P.E.: Even the Rich Can Make Themselves Poor. A Critical Examination of IV Methods in Marketing Applications. Marketing Science 33, 655–672 (2014)
11. Vaver, J. and Koehler, J.: Measuring Ad Effectiveness Using Geo Experiments, static.googleusercontent.com/media/research.google.com/en//pubs/archive/38355.pdf (Accessed: 01.12.2016)

12. Angrist, J.D., Pischke, J.-S.: Mostly harmless econometrics. An empiricist's companion. Princeton Univ. Press, Princeton, NJ (2009)
13. Abraham, M.: The Off-Line Impact of Online Ads. Harvard Business Review 86, 28 (2008)

Architecture and Evaluation Design of a Prototypical Serious Game for Business Information Visualization

Christian Karl Grund[1], Michael Schelkle[1], Max Hurm[1]

[1]University of Augsburg, Faculty of Business and Economics, Augsburg, Germany
{christian.grund,michael.schelkle}@wiwi.uni-augsburg.de

Abstract. Poorly visualized business reports may lead to wrong decisions caused by incomprehensible or misleading data. However, many companies still do not strive for adequate business information visualization (BIV), which may be due to a lack of knowledge about how to achieve it. To support managers in avoiding the pitfalls of incomprehensible reports, we are currently developing a serious game that helps players to learn about guidelines for adequate BIV. In this so-called "Dashboard Tournament", players compete across several minigames that address specific BIV guidelines. The aim of this paper is to provide an understanding of the prototype's architecture and to propose an experimental design for its evaluation. Researchers and practitioners may hence increase their understanding of how to design and evaluate serious games in the domain of business and information systems engineering.

Keywords: Serious Games, Business Information Visualization, Game-based Learning, Prototype.

1 Introduction

Poorly visualized business reports may lead to wrong decisions due to incomprehensible or misleading data [1]. Despite these threats, many companies still do not strive for proper business information visualization (BIV) [2]. One explanation for this is the lack of knowledge about adequate BIV practices and guidelines [3]. Experiential learning might be a way to sustainably increase this knowledge and therefore improve the way reports are designed [4]. Serious games are one form of experiential learning that has been used for decades to successfully convey business-related content by engaging players [5]. However, despite the plethora of different serious games described in literature, BIV has thus far not been a dedicated aspect of them [6, 7]. To fill this gap, we are developing a serious game called "Dashboard Tournament" that aims to increase BIV capabilities among players by letting them compete across several minigames [7]. Each minigame confronts players with insufficient BIV like pie charts, traffic lights, or crowded tables in reports. After describing the concept of the game in prior research [7], we aim to present its architecture and propose an experimental design for its evaluation in this paper. This may provide researches and practitioners with insights about how to develop and evaluate serious games in the domain of management reporting.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Grund, C.K.; Schelkle, M.; Hurm, M. (2017): Architecture and Evaluation Design of a Prototypical Serious Game for Business Information Visualization, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1271-1274

2 Theoretical Background and Development Method

Since serious games are concerned with improving player capabilities as well as providing an entertaining experience [8], both learning and motivation theories are used in literature to explain the benefits of serious games [9]. For instance, they are often described as a form of experiential learning [4]. To explain the motivational effects of our game, we draw on self-determination theory [10]. According to this theory, video games in general foster intrinsic motivation by enabling perceived competence, autonomy, and relatedness [11]. We hence also expect to increase intrinsic motivation with our game by satisfying these needs. Perceived competence may be fostered by players succeeding in the different minigames and earning points for doing so. Relatedness may be achieved by letting players compete in the same room and using leaderboards that allow comparisons with other players. Last, a sense of autonomy may be achieved by players being able to choose their own approaches of how to succeed in the minigames. To develop the Dashboard Tournament, we employ the human-centred design process [7]. In the following, we describe the architecture of an evolutionary prototype that resulted from the first iteration of this development process.

3 Architecture of the Dashboard Tournament

The prototype of the Dashboard Tournament currently features a singleplayer mode that comprises four minigames [7]. To implement the prototype, we used the game engine Unity with C# as the programming language. An overview of the game's architecture is provided in Figure 1.

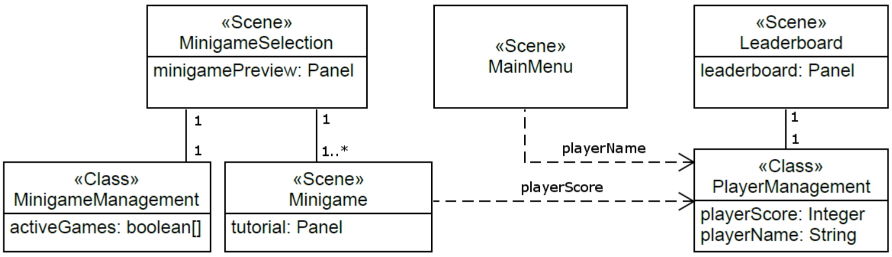


Figure 1. Architecture of the Dashboard Tournament

The game comprises different scenes (i.e., screens that players will access during the course of the game), classes that store the data necessary for the scenes to operate as well as several panels (i.e., graphical elements inside the scenes). First, players enter the main menu ("MainMenu") where they can enter their nicknames, which will be stored in the "PlayerManagement" class. Afterwards, a scene where the next minigame gets selected at random ("MinigameSelection") is shown. The different minigames are represented as "minigamePreview" panels in this scene. After the minigame that has to be played is selected, players access the respective scene for that minigame ("Minigame"). Each minigame features a tutorial panel that provides players with information regarding the objective of the current minigame and how to play it. When the minigame is finished, scores are saved in the "PlayerManagement" class and players enter a scene for displaying leaderboards ("Leaderboard"). Here they will find their score on a leaderboard panel. Afterwards, they return to the scene "MinigameSelection" as long as there are minigames left to be played. This information is stored in the "MinigameManagement" class. Although gameplay data is currently only available at runtime, a log file is going to be available on the server in later versions of the game for analysis purposes. Due to the prototype's component-based architecture, minigames may be added or removed in future iterations of the development process. In addition, multiplayer functionality will be added by defining one instance of the game as a host that selects minigames and keeps all clients synchronized.

4 Evaluation of the Prototype and Conclusion

To evaluate the game after its development will be finished (i.e., multiplayer functionality is added), we plan to conduct a laboratory experiment using a multivariate 1x3 between-group design (see Table 1). Power analysis revealed that for statistically significant results ($d = 0.8; \alpha = 0.05; 1 - \beta = 0.95$), each group should consist of 35 participants who are randomly assigned from a pool of students in business and economics programs (i.e., prospective managers and report designers).

Table 1. Experimental Design of the Evaluation

Group	Pretest	Treatment	Post-Experience	Posttest
1	Suggestions	Competition	Intrinsic Motivation	Suggestions
2	Suggestions	Singleplayer	Intrinsic Motivation	Suggestions
3	Suggestions	Presentation	Intrinsic Motivation	Suggestions

The treatments differ in how they aim to increase BIV capabilities. In the first treatment, participants play the Dashboard Tournament in a competition. The second treatment uses a modified version of the game, where there is no competition at all. This condition is used to isolate the effect of providing a competition: If the singleplayer version leads to the same benefits, the game may be easier to use in practice, since it would not require several managers to attend the same session. Last, there is a treatment with only a presentation about BIV guidelines, serving as a control group. To assess the motivational benefits of the game, we conduct post-experience questionnaires regarding perceived competence, autonomy, and relatedness as well as intrinsic motivation of participants by using the intrinsic motivation inventory [12]. To assess learning outcomes, pre- and posttests are going to address participants' BIV capabilities. For this purpose, participants are provided with different examples of business reports and are requested to suggest improvements. The provided reports suffer from inadequate BIV that is addressed by the guidelines covered in the different treatments. We can hence check whether improvements suggested by participants comply with the BIV guidelines. The pretests also help in determining prior knowledge of participants (e.g., courses or practical experience).

By comparing the post-experience questionnaires of all treatments, we may investigate whether playing the game leads to increased motivation compared to hearing a presentation. To examine the effect of setting up a competition, we may look for differences in motivation between providing a competition between players and simply playing the minigames (first and second treatment). We may also compare the learning outcomes in all treatments to see whether participants who play the game actually show increased BIV capabilities compared to participants only hearing a presentation. Last, we intend to examine correlations between motivation and learning outcomes.

In summary, this evaluation may show that the Dashboard Tournament leads to increased motivation as well as increased learning outcomes. This may encourage both researchers and practitioners to consider using serious games in the domain of management reporting. Since our approach appears to be the first serious game about BIV guidelines [6, 7], we intend to investigate its usage in this domain in future research. Especially the importance and effects of competition can be examined in further studies. By describing an architecture as well as proposing an evaluation of our game, we also aim to support building and evaluating these games.

References

1. Ware, C.: Information Visualization: Perception for Design. Elsevier (2012)
2. Al-Kassab, J., Schiuma, G., Ouertani, M.Z., Neely, A.: Information Visualization to Support Management Decisions. International Journal of Information Technology and Decision Making 13, 407–428 (2014)
3. Few, S.: Show me the Numbers. Designing Tables and Graphs to Enlighten. Analytics Press, Burlingame, Calif. (2012)
4. Kolb, D.A.: Experiential learning: Experience as the source of learning and development. Prentice-Hall Englewood Cliffs, NJ (1984)
5. Faria, A.J., Hutchinson, D., Wellington, W.J., Gold, S.: Developments in Business Gaming. A Review of the Past 40 Years. Simulation & Gaming 40, 464–487 (2009)
6. Grund, C.K., Meier, M.C.: Towards Game-based Management Decision Support: Using Serious Games to Improve the Decision Process. In: Proceedings of the Multikonferenz Wirtschaftsinformatik (MKWI) 2016, pp. 155–166 (2016)
7. Grund, C.K., Schelkle, M.: Developing a Serious Game for Business Information Visualization. In: Proceedings of the 22nd Americas Conference on Information Systems (AMCIS) (2016)
8. Abt, C.C.: Serious Games. University Press of America, Lanham, MD (1987)
9. Grund, C.K.: How Games and Game Elements Facilitate Learning and Motivation: A Literature Review. In: Informatik 2015, pp. 1279–1293. Ges. für Informatik, Bonn (2015)
10. Deci, E.L., Ryan, R.M.: Intrinsic Motivation and Self-Determination in Human Behavior. Plenum Press, New York (1985)
11. Ryan, R.M., Rigby, C.S., Przybylski, A.: The Motivational Pull of Video Games: A Self-Determination Theory Approach. Motivation and Emotion 30, 344–360 (2006)
12. Ryan, R.M., Deci, E.L.: Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. American Psychologist 55, 68 (2000)

One Plug at a Time – Designing a Peer-to-Peer Sharing Service for Charging Electric Vehicles

Jan H. Betzing[1], Moritz von Hoffen[1], Florian Plenter[1],
Friedrich Chasin[1], Martin Matzner[2], and Jörg Becker[1]

[1]University of Münster, European Research Center for Information Systems (ERCIS), Münster, Germany
{jan.betzing,moritz.von.hoffen,florian.plenter,friedrich.chasin, joerg.becker}@ercis.uni-muenster.de
[2]University of Erlangen-Nürnberg, Digital Industrial Service Systems, Nuremberg, Germany
martin.matzner@fau.de

Abstract. The widespread diffusion of electric vehicles (EVs) suffers from the lack of a well-developed public charging infrastructure, which currently is uneconomical to develop for investors. Many owners of EVs have private charging stations at their premises, which yield unlockable potential due to high idling times. In line with sharing platforms for other goods like accommodations and cars, we present the design, prototypical implementation, and evaluation setting of *CrowdStrom*, a peer-to-peer sharing service for charging EVs that networks individuals, their charging stations, and charging service customers.

Keywords: Electric Vehicle, Charging Infrastructure, Peer-to-Peer, Sharing, Design Science Research

1 Infrastructures for Electric Vehicle Charging

Clean, carbon-neutral transportation, fueled by electricity generated from renewable sources has been on governments' agendas for years. Central to these ambitious development plans is the potentially widespread diffusion of electric vehicles (EVs) [1]. While the German government, which is pursuing a target of one million registered EVs in Germany by 2020 [2], provides financial incentives for EV-buyers, registrations remain low with little more than 50,000 EVs registered by the end of 2015 [2]. In a circular pattern, customer adoption of EVs has been hampered by the lack of a well-developed public charging infrastructure [1], and the resulting limited demand for EVs deters investors from developing these infrastructures. Many owners of EVs have private charging stations at their premises, which yield unlockable potential due to high idling times. Peer-to-Peer (P2P) sharing has been proposed as way to increase the availability and use of existing resources by granting public access to individually owned resources [3]. P2P sharing facilitates IT-enabled economic transactions between individuals, where only temporary access to a physical resource is granted by a resource owner to others who need it in exchange for a monetary consideration [4]. The shift in

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Betzing, J. H.; von Hoffen, M.; Plenter, F.; Chasin, F.; Matzner, M.; Becker, J. (2017): One Plug at a Time - Designing a Peer-to-Peer Sharing Service for Charging Electric Vehicles, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1275-1278

consumer attitude from preferring ownership of a resource to preferring access to it has fueled the rise of IT-enabled P2P sharing services in domains such as transportation (e.g. *BlaBlaCar*) and accommodation (e.g. *Airbnb*) [3]. The central component of such services is a Web platform that brings together peer-providers and peer-consumers and that provides a framework for handling business transactions [5].

Against this background, we adopt the P2P sharing paradigm to enable individuals and businesses to share their private charging stations with other EV-drivers.

The remainder of the paper is structured as follows: section two sketches the research approach and the activities conducted in the course of the CrowdStrom project. Section three introduces the business model and the CrowdStrom platform, while the fourth section closes with the discussion of a comprehensive evaluation setting.

2 Research Approach

We adopt the Design Science Research (DSR) paradigm [6] to develop a business model and an instantiation of a P2P service for sharing EV-charging stations within a joint industry-and-academia consortium. In line with research guidelines on DSR [6], our work follows the cyclic process of the DSR Methodology (Figure 1).

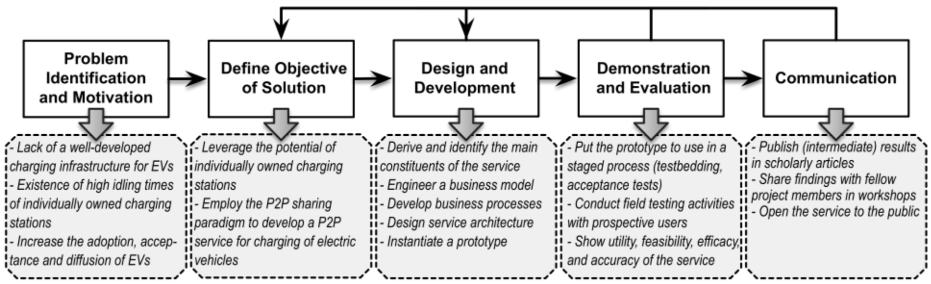


Figure 1. Overview of the DSR phases adopted for the CrowdStrom project

After identifying the mentioned societal problem, we further developed and communicated the project's objectives in multiple design cycles [7]. During the design and development phase, we addressed different issues of the EV-charging challenge ranging from customer-related (acceptance, willingness-to-pay), provider-related (incentives, organizational development), and legal aspects (energy provision, taxation) to service design (procedural/ technical design principles) [4]. We will subsequently conduct a comprehensive evaluation to advance the prototype into a mature solution.

3 The CrowdStrom Business Model and Prototype

The CrowdStrom service networks individuals and businesses that provide charging stations (peer-providers), and EV-drivers that require charging (peer-customers). Additionally, a service operator acts as intermediary that runs the central Web platform, handles transactions, and provides support services in return for a share of the revenue. The roles of the peer-provider and peer-customer are not mutually exclusive, i.e., users may both provide charging stations and use the ones of other users. Peer-providers connect their charging stations to the platform and define hourly prices according to location and power output. While the service is not intended to generate large revenue for peer-providers, it reduces the total cost of ownership of their already existing and individually bought charging stations. Furthermore, we identified a strong solidarity within EV-drivers and societal aspects to positive influence peers to become providers [7]. Peer-customers register at the service and then use the Web platform to search for and possibly reserve charging stations nearby. Additional filters for distance, price, charge rates [kW/h], plug types, green power, and customer ratings may also be applied to the search process. Peer-customers authenticate themselves at the charging station by means of RFID cards, text messages or the Web front-end. Charging sessions are recorded and fed back to the platform, so both providers and customers can have dashboards that transparently reflect their service usage (charging sessions, total energy transferred, current balance, invoices) and to facilitate an automated billing process that settles balances monthly.

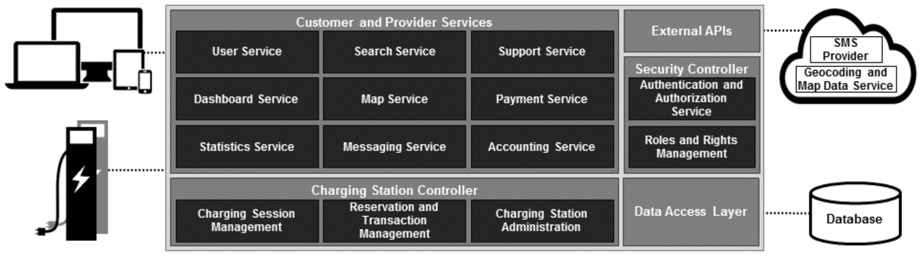


Figure 2. CrowdStrom Platform Components and Services Architecture

Figure 2 depicts the implemented solution design of the prototypical CrowdStrom platform. We leverage state-of-the-art open-source technologies (e.g. *AngularJS, Materialize*) for the responsive Web front-end and the server back-end (e.g. *Java EE, MySQL*). The charging station controller handles communication between charging stations and the platform, using the de facto standard *OCPP*. We rely on third-party providers for geocoding, cartographic material, and SMS handling. A central security service manages user authentication and authorization based on privileges and roles. Depending on the user's role, a range of functions and services is offered.

4 Evaluation and Outlook

VENABLE et al. define four evaluation strategies for DSR projects [8], from which we selected the *Human Risk & Effectiveness* strategy, as it focusses on socio-technical aspects of the service and on its real-world application. In earlier design cycles, we iteratively improved CrowdStrom's business model, processes, and service delivery mechanisms by conducting artificial formative evaluations in form of expert interviews, simulated experiments, and workshops with domain experts and technicians from our partnering local utility. For the final evaluation, we will conduct a naturalistic summative evaluation [8] in which we rigorously evaluate the validity and effectiveness of the prototype and its underlying design to ensure their utility in the real-world environment they were designed for. We will conduct two field tests for which we invite potential users to become peer-providers or peer-customers to gain insights regarding both roles.

For the first field test, we provide participants with a charging station to set up at their premises and connect to CrowdStrom in order to experience being a peer-provider for one day. During the course of the test, project members will mimic customers and perform charging sessions. Afterwards, we conduct semi-structured interviews with the participants to understand their experiences regarding the setup process, functionalities of the prototype, financial prospects, and interaction with test customers.

For the second field test, we scheduled 30 participants to perform a one-hour long test drive in an EV that entails comprehensive use of the CrowdStrom platform, i.e., search for a charging station, reserve it, visit it, and conduct a charging session. To assess the perceived ease-of-use of the solution, participants will answer semi-structured interviews. In addition, their interaction with the prototype will be screen-captured and monitored. Since both field tests will be very resource intensive (fleet of EVs, charging stations, personnel) the number of participants, especially on the provider side, will be rather small. Nevertheless, by closely monitoring and surveying the participants, we are sure to receive fruitful evaluation results that will help to improve CrowdStrom, reduce uncertainties about social and use issues, validate the ability of the design to solve the initially identified problems, and inform the launch of the platform.

Acknowledgements

This paper has been written in context of the research project CrowdStrom (www.crowdstrom.de). The project is funded by the German Federal Ministry of Education and Research (BMBF), promotion sign 01FK13019E.

References

1. Rezvani, Z., Jansson, J., Bodin, J.: Advances in consumer electric vehicle adoption research: A review and research agenda. Transp. Res. Part D Transp. Environ. 34, 122–136 (2015)
2. German National Platform for Electric Mobility: Wegweiser Elektromobilität. (2016)
3. Belk, R.: You Are What You Can Access: Sharing and Collaborative Consumption Online. J. Bus. Res. 67, 1595–1600 (2014)
4. Matzner, M., Chasin, F., von Hoffen, M., Plenter, F.: Designing a Peer-to-Peer Sharing Service as Fuel for the Development of the Electric Vehicle Charging Infrastructure. In: HICSS 2016 Proceedings. pp. 1587–1595 (2016)
5. Andersson, M., Hjalmarsson, A., Avital, M.: Peer-to-Peer Service Sharing Platforms: Driving Share and Share Alike on a Mass-Scale. In: ICIS 2013 Proceedings (2013)
6. Peffers, K., Tuunanen, T., Rothenberger, M., Chatterjee, S.: A Design Science Research Methodology for IS Research. J. Manag. Inf. Syst. 24, 45–77 (2007)
7. Matzner, M., Chasin, F., Todenhöfer, L.: To Share or not to Share. In: ECIS 2015 Proceedings (2015)
8. Venable, J., Pries-Heje, J., Baskerville, R.: FEDS: a Framework for Evaluation in Design Science Research. Eur. J. Inf. Syst. 25, 77–89 (2016)

Feldexperiment zur Wirksamkeit von konkretem vs. abstraktem Eco-Driving Feedback

André Dahlinger[1], Felix Wortmann[1], Verena Tiefenbeck[2], Ben Ryder[2], Bernhard Gahr[1]

[1] Universität St. Gallen, ITEM, St. Gallen, Schweiz
{andre.dahlinger, felix.wortmann, bernhard.gahr}@unisg.ch
[2] ETH Zürich, MTEC, Zürich, Schweiz
{vtiefenbeck, bryder}@ethz.ch

Abstract. Eco-Driving Feedback Informationssysteme haben das Potenzial, auf sehr rasche und effiziente Weise zur Reduktion von Spritverbrauch und CO2 im Strassentransportwesen beizutragen. Vor allem neuere vernetzte Fahrzeuge bieten durch die zunehmende Digitalisierung der Cockpits viel Freiraum zur Gestaltung und Untersuchung von Eco-Driving Feedback. Neben klassischen „Spritverbrauchs-Tachos" findet man heutzutage auch eher abstrakte Formen der Informationsdarbietung, z.B. animierte Pflanzen, die in Abhängigkeit vom Spritverbrauch wachsen oder schrumpfen. Diese Möglichkeiten werfen wichtige Fragen auf: Welche Art der Informationsdarbietung ändert das Fahrverhalten kurzfristig und langfristig effektiver? Welche Fahrverhaltensweisen (z.B. Bremsverhalten) werden in welchem Masse beeinflusst? Hängen eventuelle Effekte von bestimmten Moderatoren ab? Um diese Fragen zu untersuchen, wurde ein Eco-Driving Feedback System entwickelt, welches den Spritverbrauch des Fahrzeugs auf klassisch „konkrete" Art einerseits und auf eher abstrakte Art andererseits anzeigt. Die Auswirkungen auf Fahrverhalten und Spritverbrauch sollen in einem Feldexperiment unter Alltagsbedingungen mit 72 Pannenservice-Fahrern untersucht werden.

Keywords: Eco-Driving, Persuasive Technology, Feedback.

1 Einleitung und Forschungsfragen

Die Möglichkeit, Personen und Güter über Strassen zu transportieren, war und ist ein entscheidender Faktor für Aufbau und Wahrung von Wohlstand und individueller Mobilität unserer heutigen Gesellschaft, geht jedoch auch mit substantiellen negativen Externalitäten einher: 2010 gingen nach OECD Angaben 17% des weltweiten CO_2-Ausstosses auf den Strassenverkehr zurück [1]. Als Reaktion darauf verschärfen viele Staaten ihre Regularien zum CO_2-Ausstoss für Fahrzeuge. Entsprechende Verbesserungen der Technik durch die Fahrzeughersteller sind jedoch sehr teuer und langwierig [2]. Als kosteneffizienter und schneller einsetzbar gilt es daher, das Fahrverhalten der Fahrzeugnutzer selbst zu adressieren. Eine spritsparende Fahrweise kann den Verbrauch und CO_2-Ausstoss schlagartig um bis zu 30% reduzieren [3]. Um

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Dahlinger, A.; Wortmann, F.; Tiefenbeck, V.; Ryder, B.; Gahr, B. (2017): Feldexperiment zur Wirksamkeit von konkretem vs. abstraktem Eco-Driving Feedback, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1279-1282

Fahrzeuglenker zu einem spritsparenderen Fahrverhalten zu bewegen, hat sich in Untersuchungen zumeist Echtzeit-Feedback, also die zielgerichtete Rückmeldung über verhaltensrelevante Informationen in der Situation, als die effektivste Interventionsmethode herausgestellt [4]. Aufgrund ihrer Möglichkeiten zur Informationsverarbeitung und -Darbietung bieten sich Informationssysteme (IS) optimal zur Gestaltung von derartigem Feedback zur Verhaltensintervention an [5]. Sie sind entsprechend auch schon seit Jahrzehnten in Auto-Cockpits zu finden (Abbildung 1a), wo sich derzeit jedoch ein fundamentaler Wandel ereignet.

Digitale Instrumentenanzeigen finden zusehends Einzug in heutige Fahrzeugcockpits [6]. Fahrzeughersteller nutzen diese Möglichkeiten bereits zur Implementierung unterschiedlicher Eco-Driving-Feedback-Informationssysteme (EDFIS). Besonders auffällig dabei ist, dass sich vor allem der Grad der Abstraktheit der dargebotenen Information unterscheiden kann. Einerseits finden sich EDFIS mit hohem Detailgrad und einer Fülle an Informationen zu Eco-Driving-relevanten Faktoren (z.B. Brems- und Beschleunigungsverhalten, siehe Abbildung 1b). Das andere Extrem bilden EDFIS, welche statt konkreter Zahlen abstrakte Repräsentationen der relevanten Information darbieten. Statt dem Spritverbrauch in Litern pro 100 Kilometern (l/100km) sieht der Nutzer dann zum Beispiel eine Pflanze, die gedeiht oder verwelkt in Abhängigkeit vom Spritverbrauch (Abbildung 1c).

Obwohl Fahrzeughersteller EDFIS seit Jahrzehnten anbieten, finden sich in der Literatur zu EDFIS folgende Forschungslücken: (1) Es finden sich kaum belastbare Untersuchungen zur Wirkung von EDFIS auf Spritverbrauch mit sauberen Forschungsdesigns [7]. (2) Ebenso fehlen trotz aktueller Relevanz Forschungsergebnisse zur Frage, welche Rolle unterschiedliche Designfaktoren für die Wirksamkeit von Feedback-IS spielen [8]. Forschung zur Construal-Level Theory deutet darauf hin, dass abstraktes zusätzliches Potential zur Förderung von spritsparendem Verhaltens bietet [9]. Um diese Forschungslücken zu schliessen, soll im Folgenden ein Prototyp eines EDFIS vorgestellt werden, welcher Feedback in „konkreter“ und in „abstrakter“ Art darstellen kann und mit dessen Hilfe in einem gross angelegten Feldexperiment folgende Forschungsfragen verfolgt werden: (1) Reduziert ein EDFIS den Spritverbrauch bei Pannenhilfefahrern? (2) Ist abstraktes Feedback wirksamer als konkretes?

(a) (b) (c)

Abbildung 1. Unterschiedliche EDFIS: (a) klassische Spritverbrauchsanzeige im BMW 7, Bj. 1982; (b) modernes EDFIS in Jaguar Land Rover Modellen mit detaillierter konkreter Information; (c) modernes abstraktes EDFIS der Ford SmartGauge mit animierter Pflanze

2 Beschreibung des EDFIS Prototyps und der Stimuli

Der verwendete EDFIS-Prototyp bestehet aus einem On-Board-Diagnostics (OBD2) Dongle, einem Smartphone mit App und einem Backend-Server. Der OBD2-Dongle wurde derart konfiguriert, dass er Fahrzeug-CAN-Bus-Daten mit einer Frequenz von 1-30Hz auslesen kann. Diese Daten werden über Bluetooth an das Smartphone gesendet, welches relevante Informationen über eine eigens entwickelte App dem Fahrer visuell darbietet und sie gleichzeitig über eine GSM-Verbindung, angereichert mit Smartphonedaten (z.B. GPS-Position, Uhrzeit), an ein Backend schickt, wo alle Daten verarbeitet und gespeichert werden.

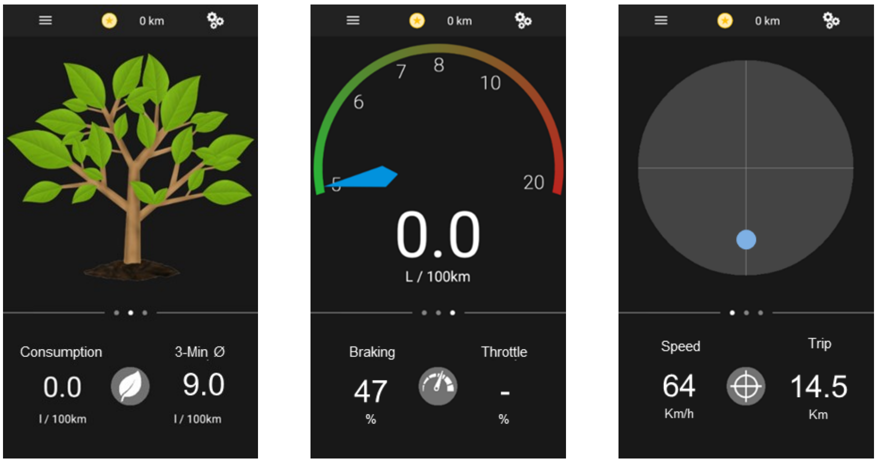
Um die dargestellten Forschungsziele zu erreichen, wurden drei Feedback-Screens entwickelt, um die jeweilige Auswirkung auf das Fahrverhalten im Feldexperiment gegeneinander testen zu können. (1) Der „Eco-Baum"-Screen spiegelt den Spritverbrauch auf abstrakte Art wieder, indem ein Baum angezeigt wird, welcher in Abhängigkeit vom Spritverbrauch wächst bzw. schrumpft (Abbildung 2a). Als Referenzwert für den Baum wurde dabei der gleitende Mittelwert des Spritverbrauchs in l/100km über die letzten 3 Minuten herangezogen, um ein extremes Schwanken im Wachstum des Baums zu verhindern. Der kleinste Baum repräsentiert einen Verbrauch von 18l/100km, der grösste Baum repräsentiert 7l/100km. Der momentane Spritverbrauch wurde ebenfalls angezeigt, um die Nachvollziehbarkeit des Screens zu erhöhen. (2) Der „Eco-Tacho"-Screen zeigt den momentanen Spritverbrauch in Echtzeit in Form einer Tachonadel an (Abbildung 2b). Darüber hinaus werden Bremsung in Prozent (-1g=100%) und Gaspedalposition in Prozent (voll durchgetreten=100%) angezeigt. (3) Der „G-Radar"-Screen (Abbildung 2c) zeigt den Vektor aus longitudinaler (vertikale Achse) und lateraler (horizontale Achse) Beschleunigung in g in Form eines blauen Punktes auf einem Kreis an (Kreisrand entspricht Beschleunigung von 1g). Zusätzlich wurden Triplänge und momentane Geschwindigkeit angezeigt. Der G-Radar dient als Kontrollbedingung und sollte die Systemnutzung aufrechterhalten, ohne einen Einfluss auf Eco-Driving auszuüben.

3 Beschreibung des Feldexperiments

Für das Feldexperiment wurden 72 Pannenhilfe-Fahrzeuge mit unserem EDFIS ausgestattet. Die teilnehmenden Fahrer wurden per interner E-Mail rekrutiert. 72 von 92 angeschriebenen Fahrern nahmen freiwillig am Experiment teil.

Das Feldexperiment begann mit einer zwei-wöchigen Baseline-Phase. Während dieser sahen alle Fahrer den G-Radar. In der anschliessenden Interventionsphase wurden die Fahrer randomisiert einer von drei Gruppen zugewiesen: (1) die Kontrollgruppe sah weiterhin den G-Radar (N=25), (2) die Gruppe mit konkretem Feedback wechselte automatisch auf den Eco-Tacho (N=24) und (3) die Gruppe mit abstraktem Feedback wechselte automatisch auf den Eco-Baum (N=23).

In der Analyse sollen künftig die Spritverbräuche der drei Gruppen verglichen werden, um so Rückschlüsse auf die Wirksamkeit der unterschiedlichen EDFIS auf den Spritverbrauch im Sinne der dargestellten Forschungsfragen zu erlangen.



(a) "Eco-Baum" (b) "Eco-Tacho" (c) "G-Radar"

Abbildung 2. Die Feedback Screens für die (a) Gruppe mit abstraktem Eco-Feedback; (b) Gruppe mit konkretem Eco-Feedback; (c) Kontrollgruppe und Baselinephase

Acknowledgements

Diese Arbeit wird unterstützt vom Bosch IoT-Lab, Universität St. Gallen, CH.

4 References

1. ITF: Transport Green House Gas Emissions - Country Data 2010. (2010).
2. GFEI: Fuel Economy State of the World 2014. (2014).
3. Barić, D., Zovak, G., Periša, M.: Effects of Eco-Drive Education on the Reduction of Fuel Consumption and CO2 Emissions. PROMET (Traffic Transp. 25, 265–272 (2013).
4. Barkenbus, J.N.: Eco-driving: An overlooked climate change initiative. Energy Policy. 38, 762–769 (2010).
5. Watson, R.T., Boudreau, M.-C., Chen, A.J.: Information systems and environmentally sustainable development: energy informatics and new directions for the is community. MIS Q. 34, 23–38 (2010).
6. GSMA: 2025 Every Car Connected. (2012).
7. Tulusan, J., Staake, T., Fleisch, E.: Providing eco-driving feedback to corporate car drivers: what impact does a smartphone application have on their fuel efficiency? UbiComp. 1–4 (2012).
8. Karlin, B., Zinger, J.F., Ford, R.: The effects of feedback on energy conservation: A meta-analysis. Psychol. Bull. 141, 1205–1227 (2015).
9. Dahlinger, A., Wortmann, F.: Fostering Pro-Environmental Behavior with Green Consumer IS: The Effects of IS-Induced Construal and General IS Usage Motivations. In: ECIS. , Istanbul (2016).

Towards a Software Prototype Supporting Automatic Recognition of Sketched Business Process Models

Manuel Zapp, Peter Fettke and Peter Loos

Institute for Information Systems at the German Research Center for Artificial Intelligence (DFKI GmbH) and Saarland University, Saarbrücken, Germany
{manuel.zapp,peter.fettke,peter.loos}@iwi.dfki.de

Abstract. The paper at hand presents a prototype that implements an approach to recognize the structure of business process models sketched on paper and whiteboards to create digital versions of those models. We explain the different steps from common photos of business process models sketched on paper or whiteboards to their digital version. We modified existing approaches from sketch recognition to fulfill the needs of sketched business process model recognition. Therefore, a dataset was generated that in turn is used to train different classifiers for different shapes within business process models.

Keywords: sketch recognition, BPM, computer vision, machine learning.

1 Introduction

In Business Process Modeling there has always been an interest in building tools assisting the modeler in creating business process models. Such tools aim at being user-friendly and intuitive. Nevertheless, pen and paper is still the predominant choice, especially in the early design phase of business processes. With no doubt using pen and paper is one of the most intuitive ways of creating business process models. Without manual work, these models can only be stored as pictures or archived on paper and not be transferred into existing modelling systems. To bridge this gap, we introduce a software system that fully automatically transfers sketched business processes into a digital format that can be interpreted by common modeling tools. This prototype is developed within the research project INDIGO in cooperation with a leading Business Process Management software provider, which has multiple customers that have expressed a strong interest in the functionalities of the prototype. Although we use the modeling convention Event-driven Process Chain (EPC) as a sample, our approach is generic and can be applied to other modelling conventions.

In our work, we want to implement an offline approach that is capable of constructing digital business process models from sketched ones on whiteboards or paper. Please note that this is work in progress. In this paper we concentrate on the structural recognition of business process models without solving the Hand Writing Recognition problem for labels that are usually present in sketched process models.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Zapp, M.; Fettke, P.; Loos, P. (2017): Towards a Software Prototype Supporting Automatic Recognition of Sketched Business Process Models, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1283-1286

2 Related Work

In general, two approaches in sketch recognition can be distinguished – offline approaches and online approaches. In online approaches the process of sketching is observed, whereby the temporal order information of strokes is leveraged to reduce the search space [1]. Offline approaches are applied to images of finished sketches only and thus are much more challenging.

A lot of work has been done in the area of online sketch recognition, especially in the field of flowchart recognition [2–4], that is by means related to business process recognition. There is only one work available in that area that deals with an offline approach [1] and we refer to it.

To the best of our knowledge we are pioneers on the field of sketched business process digitalization. There exist some works that use post-it®s and create their own modeling conventions that are designed and simplified for later digitalization.[1] The major drawbacks of these systems are the modeling conventions that first have to be learnt by all persons involved as well as the limitations of the modeling conventions and the additional hardware that is needed in form of post-it®s.

3 Our Approach and Its Implementation

3.1 Overview

The prototype is designed as a web application that provides interfaces to prepare data in form of photos taken from sketched business process models to train the software system itself. Additionally, an interface is provided that allows a user to upload images of sketched process models and responds with a representation of the recognized model. Components that actually incorporate image processing and computer vision functionalities are built in C++, due to performance reasons and better handling of the OpenCV[2] library that is used. In the following the necessary steps for an automated recognition are explained.

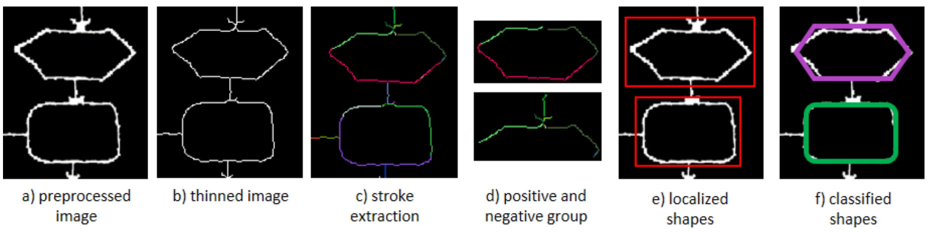


Figure 1. Different steps of processing

[1] http://www.symbioworld.com/de/produkt/produkt-portfolio/workshop-methode.html
[2] http://opencv.org/

3.2 Preprocessing and Stroke Extraction

In our work we target an end-to-end scenario. As we do not want to have unrealistic expectations at the initial quality of the images that should be processed, we allow the images to come from common hardware like smartphone cameras, but we demand sufficiently high contrast between the color of the used pen and the color of the surface the sketch is drawn on. As a consequence, some preprocessing to the initial image is necessary. We apply common filters from Computer Vision to get rid of artifacts in the image. The result of this process is a clean image where only white lines are displayed on a black background as shown in Figure 1a).

Since we follow an offline approach there is no structural information about strokes available. Based on the approach presented in [5] we reconstruct strokes from pixel data and adapt it to our needs. The authors propose to apply a Morphological Thinning algorithm [6] to the preprocessed image, so that all lines are reduced to a one-pixel width as shown in Figure 1b). The extracted strokes are sets of single pixels from the thinned image. Since these strokes may span over several different shapes, which is undesirable for our work we introduce a splitting algorithm. We trace the extracted strokes and at every *n-th* point and perform a principal component analysis on the last and the next *k* pixels of the actual stroke. The principal component analysis gives us the orientation of the last *k* and the next *k* pixels. If the angle between them exceeds an experimentally learnt threshold of 70 degrees, the stroke is split at that point. So all strokes extracted belong solely to one single shape in the process model as shown in Figure 1c).

3.3 Localization

To actually classify shapes into the right category we need to know which strokes form a shape. Common approaches extract rectangular regions from the image and test each against a classifier. Usually shapes contain text that leads to different appearances within the same type of shape. Thus we follow a technique presented in [1] that is called shapeness estimation. The authors propose a feature to construct a fast classifier that solely decides whether a set of strokes forms a shape or not. A group of strokes is a set of strokes that fulfils predefined constraints like spatial ones. The authors build the groups by breadth-first-search. We adopt their work for classical flow-chart recognition to also be able to detect sketched EPCs. Therefore, a dataset of 108 sketched EPCs was created by 36 people with expertise in business process modeling, all sketching 3 different EPCs including 24 shapes on average. The shapes were then extracted so we end up with nearly 2.600 labelled samples. Figure 1d) shows two stroke sets, where one forms a valid shape and, thus, got accepted by the shapeness classifier and an invalid and, thus, rejected one. In Figure 1e) the localized shapes are displayed.

3.4 Classification

The classification problem can be solved in different ways. We decided to choose the Histogram of Oriented Gradients feature since it is commonly used in sketch recognition [7]. We obtained best results in combination with Support Vector Machines. We train a classifier for each of the seven shapes (i.e. events, functions, application systems, organizational units and the three connector types) that are applied in the respective order to proposed stroke sets. If the maximum confidence exceeds a learnt threshold, the shape of the classifier with highest confidence is chosen otherwise the stroke set is dropped. Table 1 shows the accuracy measures for the overall model. We used 10-fold cross-validation to validate these results.

Table 1. Accuracy Measures

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
0.969	0.017	0.938	0.965	0.951	0.930

4 Conclusion and Discussion

We presented a prototype that implements an approach to construct common business process models based on the EPC modeling notation from sketched process models in terms of their structure. Our prototype is research in progress. At the moment, there is just a first evaluation of our prototype available. These first evaluation activities show that our approach is promising. Further evaluation is planned.

References

1. Wu, J., Wang, C., Zhang, L., Rui, Y.: Offline sketch parsing via shapeness estimation. In: IJCAI International Joint Conference on Artificial Intelligence (2015).
2. Bresler, M., Phan, T. Van, Prusa, D., Nakagawa, M., Hlavac, V.: Recognition System for On-Line Sketched Diagrams. Proc. Int. Conf. Front. Handwrit. Recognition, ICFHR. 2014–Decem, 563–568 (2014).
3. Lemaitre, A., Mouchère, H., Camillerapp, J., Coüasnon, B.: Interest of syntactic knowledge for on-line flowchart recognition. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 7423 LNCS, 89–98 (2013).
4. Carton, C., Lemaitre, A., Couasnon, B.: Fusion of statistical and structural information for flowchart recognition. Proc. Int. Conf. Doc. Anal. Recognition, ICDAR. 1210–1214 (2013).
5. Rajan, P., Hammond, T.: From Paper to Machine: Extracting Strokes from Images for use in Sketch Recognition. (2008).
6. Lü, H.E., Wang, P.S.P.: A comment on “a fast parallel algorithm for thinning digital patterns.” Commun. ACM. 29, 239–242 (1986).
7. Hu, R., Collomosse, J.: A Performance Evaluation of Gradient Field HOG Descriptor for Sketch Based Image Retrieval. (2013).

Mobiles Lernen für China – eine iterative Prototypenentwicklung

Sissy-Josefina Ernst[1], Andreas Janson[1], Matthias Söllner[1, 2]

[1] Fachgebiet Wirtschaftsinformatik, Universität Kassel, Kassel, Deutschland
{sissy.ernst,andreas.janson, soellner}@uni-kassel.de
[2] Institut für Wirtschaftsinformatik, Universität St Gallen, St Gallen, Schweiz
{soellner}@ unisg.ch

Abstract. Mobiles Lernen ermöglicht eine Integration von authentischen Lernszenarien in den Alltag und unterstützt somit den Aufbau von Problemlösungskompetenzen. Da zielgruppenspezifische Unterschiede die Nutzung von IT beeinflussen und sich Lehrkonzepte unterscheiden, muss die Zielgruppe bei der Entwicklung einer mobilen Lernanwendung (MLA) berücksichtigt werden. Dieser Beitrag zeigt eine iterative Prototypenentwicklung am Beispiel einer MLA für den Einsatz in chinesischen Berufsschulen. Auf Basis der IT-Kulturkonflikttheorie wird ein systematischer Gestaltungs- und Evaluationsprozess dargestellt. Die Ergebnisse zeigen, dass die systematische Zielgruppenanpassung den Erfolg einer MLA signifikant steigert.

Keywords: Mobiles Lernen, IT-Kulturkonflikttheorie, Zielgruppenanpassung

1 Einleitung

In China werden viele Ausbildungsberufe nicht in Form einer dualen Ausbildung mit einer engen Verzahnung zwischen Erwerb von Wissen und seiner praktischen Anwendung gelehrt, sondern in schulischen Einrichtungen unterrichtet. Insbesondere in Ausbildungsberufen, die neben theoretischen Kenntnissen auch Problemanalysefähigkeiten erfordern, fehlt es an realen Szenarien, um diese Kompetenzen aufzubauen. In Folge dieser Entwicklung erlebt China einen akuten Mangel an gut ausgebildeten Fachkräften. Mobiles Lernen (ML) kann dabei Teil einer Problemlösung sein. Es ermöglicht Lernenden Ort und Zeit von Lernaktivitäten frei zu wählen und mit Dozenten, Lernmaterial und Peers zu interagieren. In praktische Trainings eingesetzt, kann ML eine kontextuelle Wissensaneignung und direktes Feedback ermöglichen, um den Kompetenzerwerb zu unterstützen. Dennoch hat die Forschung gezeigt, dass in Kulturräumen wie China zum einen anders gelernt wird und zum anderen sich die Nutzung von Informationstechnologien (IT) unterscheidet, was eine Einführung von IT-Lösungen aus westlichen Ländern erschwert.

Ziel dieses Forschungsvorhabens ist daher die Entwicklung einer mobilen Lernanwendung (MLA), welche in der KFZ-Mechatroniker Ausbildung in China einen kontextuellen Kompetenzerwerb ermöglicht und zielgruppenspezifische Bedürfnisse

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Ernst, S.-J.; Janson, A.; Söllner, M. (2017): Mobiles Lernen für China - eine iterative Prototypenentwicklung, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1287-1290

berücksichtigt. Vorliegender Beitrag zeigt eine systematische Prototypenentwicklung auf, welche die Zielgruppe im Entwicklungsprozess berücksichtigt, um den Lernerfolg als abhängige Variable zu steigern. Vor diesem Hintergrund verweist die Forschung auf die IT-Kulturkonflikttheorie [1]. Diese besagt, dass die Rolle von Kultur in der Nutzung von IT nur bei Auftreten eines Konflikts sichtbar ist. Menschen sind sich nicht ihrer eigenen Kultur bewusst, bis es zu einem Konflikt mit einer „Gegenkultur“ kommt. Im folgenden Beitrag wird die IT-Kulturkonflikttheorie als Kerntheorie [2] verwendet, um die MLA während des Entwicklungsprozesses an die Werte der Zielgruppe anzupassen. Die Erhebung der Anforderungen basiert auf einer Analyse potentieller Konflikte, was wiederum eine gezielte Anpassung der MLA erlaubt.

2 Gestaltungsorientiertes Vorgehen

Das Vorhaben folgt einem gestaltungsorientierten Forschungsansatz [3], welcher eine iterative Prototypenentwicklung zeigt. Abbildung 1 veranschaulicht das Vorgehen.

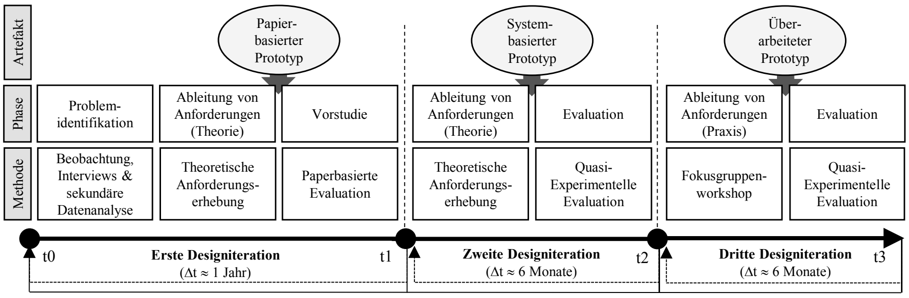


Abbildung 1. Iteratives Vorgehen zur Entwicklung der MLA

Die MLA ermöglicht kontextsensitives Lernen indem mit QR-Codes, welche an einem KFZ angebracht sind, interagiert wird. Dabei hat die MLA zwei Aufgabentypen: es müssen zum einen Autoteile gefunden und die dazugehörigen QR-Codes gescannt werden, um Basiswissen zu erwerben; zum anderen können Lösungsstrategien für Problemszenarien am KFZ durch Scannen bestimmter QR-Codes, welche jeweils für eine Fehlerursache stehen, entwickelt werden. Wissenstests ergänzen die Aufgaben.

Um die Zielgruppe im Entwicklungsprozess zu berücksichtigen, wurden in der ersten und zweiten Designiteration Anforderungen auf Basis der IT-Kulturkonflikttheorie aus der Literatur abgeleitet und somit Werte der Zielgruppe auf nationaler Ebene erfasst. Durch eine gestalterische Anpassung der MLA soll sichergestellt werden, dass die der MLA zugrunde liegenden Werte mit den IT-Werten (Vorstellungskonflikt) und generellen Wertvorstellungen der Zielgruppe (Systemkonflikt) vereinbar sind [1]. Sollte ein Konflikt zwischen den Werten auftreten, könnte dies eine nicht effektive Nutzung zur Folge haben [1, 4]. Der papierbasierte Prototyp beinhaltet hierbei Anpassungen mit Bezug zur Benutzerschnittstelle, wohingegen Konflikte hinsichtlich des Lernprozesses in dem weiterentwickelten systembasierten Prototyp instanziiert werden konnten. Die identifizierten Systemkonflikte beziehen sich auf einen lehrerzentrierten Unterrichtsstil, welcher dem nicht angeleiteten und freiwilligen Lernen des ML entgegenstehen könnte. Weiterhin sollten Nutzerdaten anonymisiert, um einen potentiellen Gesichtsverlust der Schüler bei falsch gelösten Aufgaben zu vermeiden und die Benutzerschnittstelle zielgruppenkonform gestaltet werden, da diese in verschiedenen Kulturen unterschiedliche Gestaltungselemente aufgreift. Der identifizierte Vorstellungskonflikt bezieht sich auf die IT-Werte der Zielgruppe, die ihre Smartphones primär für Unterhaltungszwecke, nicht aber zum Lernen nutzen.

Um die Werte der Zielgruppe ebenfalls auf gruppen- und individueller Ebene zu berücksichtigen, wurden in der dritten Designiteration Fokusgruppenworkshops durchgeführt. In den Workshops wurden in Gruppen weiterführende Änderungswünsche auf Grundlage einer unangepassten Basisanwendung erarbeitet.

In einem darauffolgenden Schritt wurden die jeweils erhobenen Anforderungen durch Designelemente adressiert. Um beispielsweise den zuerst genannten System- und den Vorstellungskonflikt zu adressieren, wurde ein virtueller Charakter in die MLA integriert, welcher neben Anleitungen, wie eine Aufgabe zu erfüllen ist (umgewandelt in ein Gamification-Element), direktes Feedback zu den Lernergebnissen gibt. Abbildung 2 zeigt die getesteten Prototypen aller Designiterationen.

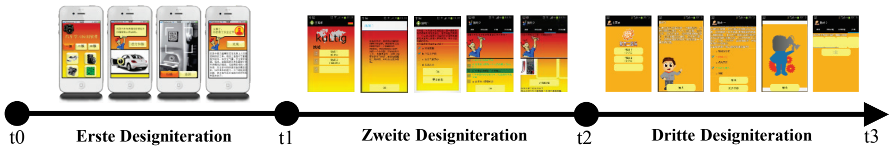


Abbildung 2. Iterative Entwicklung des Prototyps

Jede Designiteration schließt mit einer Evaluation des aktuellen Entwicklungstands. Ergebnisse sowie Feedback und Beobachtungen während der Datenerhebungen wurden zurück in das Gestaltungsvorhaben gespiegelt. So wurde beispielsweise ein Timer als Spielelement auf Grund der unkoordinierten und hektischen Nutzung der MLA revidiert und ein Männchen, das Lebensenergie bei Fehlern verliert, integriert.

3 Ergebnisse

Für die Evaluation wurden in chinesischen Berufsschulen, neben einer papierbasierten Evaluation (erste Designiteration), Quasi-Experimente (zweite und dritte Designiteration) durchgeführt. In den experimentellen Evaluationen wurde die angepasste MLA von einem Teil der Zielgruppe genutzt (Treatment) und von einem anderen Teil eine inhaltsgleiche Basisversion (Kontroll) ohne integrierte Gestaltungselemente verwendet. Während papierbasiert das Design der angepassten MLA mit dem Technology Acceptance Model [5] getestet wurde, konnte im Rahmen der Experimente neben der Technologieakzeptanz, der erworbene Lernerfolg anhand von Wissenstests erhoben werden. Das Experiment der dritten Designiteration wurde von einem Pre-Wissenstest, welcher aus 8 Wahr-Falsch Fragen bestand, eingeleitet, wobei eine Gleichverteilung zwischen Treatment- und Kontrollgruppe sichergestellt werden konnte. Die Evaluationen wurden mit einem N von 113 nach der ersten, 109 nach der zweiten und 201 Teilnehmern nach der dritten Designiteration durchgeführt. Die Teilnehmer aller Iterationen waren mehrheitlich männlich. Eine Teilnahme am Experiment war für Berufsschüler aller Jahrgangsstufen möglich. Die Evaluation des papierbasierten Prototyp (erste Iteration) zeigte eine signifikant positive Beurteilung hinsichtlich der wahrgenommenen Benutzerfreundlichkeit (Mittelwert = 5,26), wahrgenommenen Nützlichkeit (Mittelwert = 5,08) sowie der Nutzungsintention (Mittelwert = 5,12) im Vergleich zum Neutralwert der 7-stufigen Skala, was für die Tauglichkeit der MLA in China spricht. In dem Wissenstest der ersten quasi-experimentellen Evaluation (zweite Iteration), wobei Wissen anhand von Multiple-Choice Fragen erhoben wurde, konnte die Treatmentgruppe signifikant ($p < 0.001$) bessere Ergebnisse im Vergleich zur Kontrollgruppe erzielen. In der dritten Evaluation (dritte Iteration) wurde der Lernerfolg differenziert für deklaratives, prozedurales und fertigkeitsbasiertes Wissen erhoben. Für die Erhebung wurden Wahr-Falsch-Fragen sowie eine Fehlerbaumanalyse herangezogen. Es zeigten sich signifikant positive Auswirkungen der Anpassungen auf die Aneignung von prozeduralem ($p = 0.050$) und fertigkeitsbasiertem Wissen ($p < 0.001$), siehe im Detail Ernst et al. [6].

4 Fazit

Der Beitrag zeigt eine Theorie-geleitete Gestaltung und Evaluation einer MLA für Berufsschüler der KFZ-Mechatronik in China und trägt mit einer systematischen Evaluation zur gestaltungsorientierten Forschung bei [7]. Hiermit wird ein tieferes Verständnis für ML, die Anpassung an fremde Kontexte und die IT-Kulturkonflikttheorie aufgezeigt. Aus praktischer Sicht zeigt dieser Beitrag ein systematisches Vorgehen auf, um MLA zielgruppenspezifisch anzupassen.

Literatur

1. Leidner, D.E., Kayworth, T.: Review: A Review of Culture in Information Systems Research: Toward a Theory of Information Technology Culture Conflict. MIS Quarterly 30, 357–399 (2006)
2. Gregor, S.: The nature of theory in information systems 30, 611–642 (2006)
3. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. Journal of Management Information Systems 24, 45–77 (2007)
4. Koch, H., Leidner, D.E., Gonzalez, E.S.: Digitally enabling social networks: resolving IT-culture conflict. Info Systems J, 501–523 (2013)
5. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 13, 319–340 (1989)
6. Ernst, S.-J., Janson, A., Söllner, M., Leimeister, J.M.: It's about Understanding Each Other's Culture – Improving the Outcomes of Mobile Learning by Avoiding Culture Conflicts. ICIS 2016 - accepted for publication (2016)
7. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly 37, 337-A6 (2013)

Enhancing Asthma Control through IT: Design, Implementation and Planned Evaluation of the Mobile Asthma Companion

Peter Tinschert[1], Filipe Barata[2], Tobias Kowatsch[1]

University of St. Gallen, Institute of Technology Management, St. Gallen, Switzerland
{peter.tinschert,tobias.kowatsch}@unisg.ch
ETH Zurich, Department of Management, Technology and Economics, Zurich, Switzerland
fbarata@ethz.ch

Abstract. The personal and financial burden of asthma highly depends on a patient's disease self-management skill. Scalable mHealth apps, designed to empower patients, have the potential to play a crucial role in asthma disease management. However, the actual clinical efficacy of mHealth asthma apps is poorly understood due to the lack of both methodologically sound research and accessible evidence-based apps. We therefore apply design science with the goal to design, implement and evaluate a mHealth app for people with asthma, the Mobile Asthma Companion (MAC). The current prototype of MAC delivers health literacy knowledge triggered by nocturnal cough rates. We conclude by proposing a randomized controlled trial to test the efficacy of our prototype.

Keywords: Asthma Control, mHealth Apps, Asthma Disease Management

1 Introduction

Asthma, a chronic airway disease, ranks among the most prevalent noncommunicable diseases with an estimated 334 million people suffering from it globally. The yearly costs of asthma are estimated to be around 56 billion dollars in the US alone [1].

When patients are not able to control their asthma symptoms, it heavily impacts their quality of life and may even have life threatening consequences. However, successful maintenance of well-controlled asthma enables a patient to live with almost unimpaired quality of life [2]. Apart from the personal implications for patients, the degree of asthma (symptom) control profoundly affects the healthcare system from an economic point of view: A patient with uncontrolled asthma symptoms causes approximately 4.5 times the costs of a patient with well-controlled asthma, which is amplified by the fact that more than half of all asthmatics suffer from uncontrolled asthma [3].

In order to achieve asthma control, clinical guidelines emphasize the importance of an empowered patient with disease management skills [4]. Here is where mobile health applications (mHealth apps) can potentially decrease the personal and financial burden of asthma: in contrast to traditional asthma disease management programs, which

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Tinschert, P.; Barata, F.; Kowatsch, T. (2017): Enhancing Asthma Control through IT: Design, Implementation and Planned Evaluation of the Mobile Asthma Companion, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1291-1294

consist of multiple weekly face-to-face sessions guided by healthcare professionals, scalable mHealth apps can empower patients cost effectively by delivering disease management interventions to their smartphones as part of daily routine. Moreover, mHealth apps may exploit built-in smartphone sensors to objectively monitor disease symptoms and as a consequence identify windows of opportunities in which patients are particularly receptive and susceptible to respond to such interventions [5].

However, to prove the potential of mHealth asthma apps, clinical efficacy needs to be demonstrated in relation to the therapeutic gold standard (i.e. traditional asthma management programs). Systematic reviews in this field, which have indicated first promising results regarding the efficacy of mHealth asthma apps, strongly advocate for further research due to the low methodological quality of the available studies [6] and the low average quality of mHealth asthma apps, which often fail to consider clinical guidelines for asthma treatment [7].

We address this research gap by designing, implementing and evaluating our own mHealth asthma app, which is based on design science research [8]. We refer to our app as the Mobile Asthma Companion (MAC). Currently, the prototype consists of a disease management intervention focused on improving health literacy, which is triggered by, among others, automated cough detection.

The remainder of this paper is structured as follows: after introducing the applied justificatory knowledge and deriving the design requirements, we will describe the implementation of our MAC prototype. We conclude this paper by proposing an experimental study design, which will evaluate MAC's efficacy.

2 Conceptual Foundation and Design Requirements

Patients are responding particularly well to interventions when they are delivered just in time, or in other words, in the exact moment when patients demand them [5]. A state of demand is characterized, among others, by experiencing adverse health effects. For asthma, we argue that partially controlled or uncontrolled asthma indicate such a state of vulnerability. Thus, nocturnal cough rate, a valid marker for asthma control [9], could serve as an intervention trigger. We already delivered the proof of concept that cough can be monitored fully automated and accurately by means of a smartphone [10].

Requirement 1: The design artifact has to be able to monitor the nocturnal cough rate fully automated and use it to trigger just-in-time interventions.

Health literacy, or more specifically asthma education, is a key aspect of asthma disease management [4]. Research has shown that health literacy interventions are clinically efficacious in general and for asthma in specific [11].

For asthma health literacy interventions, first evidence was provided that an asthma mHealth app is able to improve asthma control significantly over the course of five weeks [12]. However, the single-arm study design and the monetary incentives for continuous study participation limit the explanatory power of the study results and the scalability of such an app. Therefore, in order to enable scalability, a user's engagement needs to be ensured without relying on financial incentives. Gamification features like point systems, badges or achievements are promising options in this regard [13].

Requirement 2: The design artifact has to deliver a health literacy intervention including educational asthma content implemented through gamification features.

3 Implementation

We implement our MAC prototype as a mobile application for Android smartphones based on the MobileCoach (MC) platform [14]. MC is an established digital health intervention platform, which provides the necessary user privacy and data security features when working with sensitive patient data. Currently, the MC based prototype offers two main functionalities: a chat and an alarm clock.

The requirement of automated nocturnal cough detection is met by expanding our prior work on cough detection [10] in order to detect cough rate overnight. The detection mode is enabled by setting an alarm, which defines a time frame from the moment the alarm is set until it rings. During this period, the detection algorithm operates in the background, automatically detecting and counting coughs.

Users will be instructed to set the MAC alarm directly before going to sleep and to place the smartphone near the bed (e.g. on a nightstand). Based on detected cough rates per night, users will be prompted to interact with MAC through push notifications. The likelihood of receiving such a notification is a function of the standardized cough rate within a user. In order to limit the burden of intervention, users can only receive up to two notifications per day. However, notifications will be sent at least twice a week to users with a particularly low cough rate who otherwise might not receive any notifications at all. Additionally, MAC is also accessible at will.

In order to address the second design requirement, MAC will interact with the user via a chat interface and provide educational material related to asthma health literacy topics in form of knowledge nuggets (e.g. video clips and quiz questions). This chat functionality is enabled by extending the current MC platform with the possibility to communicate with the user via an internet based chat client mobile application instead of a SMS based communication. Furthermore, the educational materials are fully congruent with the content of current disease management programs. The MC rule-based engine triggers the delivery of knowledge nuggets based on the nocturnal cough rate of the user. Finally, gamification features are implemented through badges which users can earn by viewing educational material and answering quiz questions correctly.

4 Study Design

We will evaluate the efficacy of our MAC prototype in a randomized control trial, the methodological gold standard for investigating efficacy. Adult asthmatics will be randomly assigned to two different groups: The control group will participate in a traditional asthma disease management program offered by local healthcare entities [4] whereas the experimental group will interact with MAC. The primary endpoint in this study is asthma control on a monthly basis. Asthma control is measured through a standardized test [15]. Study duration will be three months to account for lagged effects of the intervention on asthma control. Ideally, our study will show that MAC will perform at least as good as the control group in the measurement of the primary endpoint upon study completion (i.e. non-inferiority trial). We will consider seasonal differences in asthma by measuring asthma control additionally at study initiation. A follow-up measurement after three months will account for long-term effects.

References

1. Global Asthma Network: The Global Asthma Report 2014. Global Asthma Network, Auckland (2014)
2. Haughney, J., Price, D., Kaplan, A., Chrystyn, H., Horne, R., May, N., et al.: Achieving asthma control in practice. Respiratory Medicine 102, 1681-1693 (2008)
3. Accordini, S., Corsico, A.G., Braggion, M., Gerbase, M.W., Gislason, D., Gulsvik, A., et al.: The cost of persistent asthma in Europe: an international population-based study in adults. International archives of allergy and immunology 160, 93-101 (2013)
4. GINA: Global Strategy for Asthma Management and Prevention. GINA, Kolding (2016)
5. Nahum-Shani, I., Hekler, E.B., Spruijt-Metz, D.: Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. Health Psychology 34, 1209 (2015)
6. Marcano Belisario, J.S., Huckvale, K., Greenfield, G., Car, J., Gunn, L.H.: Smartphone and tablet self management apps for asthma. Cochrane Database Syst Rev 11 (2013)
7. Huckvale, K., Morrison, C., Ouyang, J., Ghaghda, A., Car, J.: The evolution of mobile apps for asthma. BMC medicine 13, 1-15 (2015)
8. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Quarterly 28, 75-105 (2004)
9. Marsden, P.A., Satia, I., Ibrahim, B., Woodcock, A., Yates, L., Donnelly, I., et al.: Objective Cough Frequency, Airway Inflammation and Disease Control in Asthma. Chest (2016)
10. Barata, F., Kowatsch, T., Tinschert, P., Filler, A.: Personal MobileCoach: tailoring behavioral interventions to the needs of individual participants. Proceedings of the 2016 ACM UbiComp, pp. 1089-1094. ACM, Heidelberg, Germany (2016)
11. Miller, T.A.: Health literacy and adherence to medical treatment in chronic and acute illness: A meta-analysis. Patient education and counseling (2016)
12. Cook, K.A., Modena, B.D., Simon, R.A.: Improvement in Asthma Control Using a Minimally Burdensome and Proactive Smartphone Application. The Journal of Allergy and Clinical Immunology: In Practice (2016)
13. Thiebes, S., Lins, S., Basten, D.: Gamifying Information Systems-A Synthesis of Gamification Mechanics and Dynamics. The 22nd European Conference on Information Systems (ECIS) 2014, Tel Aviv, Israel (2014)
14. Filler, A., Kowatsch, T., Haug, S., Wahle, F., Staake, T., Fleisch, E.: MobileCoach: A Novel Open Source Platform for the Design of Evidence-based, Scalable and Low-Cost Behavioral Health Interventions - Overview and Preliminary Evaluation in the Public Health Context. 14th annual Wireless Telecommunications Symposium (WTS 2015). IEEE, USA, New York (2015)
15. Nathan, R.A., Sorkness, C.A., Kosinski, M., Schatz, M., Li, J.T., Marcus, P., et al.: Development of the asthma control test: a survey for assessing asthma control. Journal of Allergy and Clinical Immunology 113, 59-65 (2004)

Live Query – Visualized Process Analysis

Dennis M. Riehle[1], Steffen Höhenberger[1], Rainer Cording[1], and Patrick Delfmann[2]

[1] University of Münster, European Research Center for Information Systems, Münster, Germany
{riehle,hoehenberger,cording}@ercis.uni-muenster.de
[2] University of Koblenz-Landau, Institute for Information Systems Research, Koblenz, Germany
delfmann@uni-koblenz.de

Abstract. Business process management (BPM) becomes continuously challenging through a steadily increasing number and even more complex processes. For enabling an effective and efficient control of business processes, (semi-)automatic approaches are necessary as a supporting means. However, these approaches are often hardly applicable in practice since they lack a broad applicability or an acceptable ease of use. This work aims to close this gap by providing an approach that supports a widely applicable, (semi-)automatic analysis of business process models and makes the analysis comprehensible using a graphical visualization.

Keywords: business process analysis, model query, visualization, ease of use

1 Motivation

The view on value creation in companies has changed during the last decades off from the function perspective to the process perspective [1], [2]. Against this background, Business Process Management (BPM) is an indispensable part of business. However, this task becomes increasingly difficult due to a steadily rising number of more and more complex processes. Thus, it is becoming harder for supervisory bodies to prevent inefficiencies, compliance violations [3] or wrong executions due to lacking overview.

Process models that conceptually represent business processes using standardized modeling languages aim to provide a remedy for this issue. Nevertheless, if the processes exhibit growing size and complexity, the models become vast and confusing, too. Hence, manual control is hardly possible, and common understanding of processes gets lost. Actually, even process models cannot reduce the complexity of real processes.

This problem can be addressed by (semi-)automatic analysis approaches [4] that support an investigation of process models at design time. In the literature, there are plenty of those. Against that, these approaches do mostly not find their way into practice. Three critical aspects mainly cause this: lacking open availability, lacking wide applicability [5], [6] and lacking ease of use [7]. While open availability is the less problematic issue to be solved in most cases, wide applicability and a sufficient ease of use are core problems of many approaches. Approaches have to be widely applicable as different modeling languages are employed in different (and sometimes even in the same) companies. Moreover, there is no common understanding of processes, e.g., in terms of granularity or terminology in many cases. Hence, the

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Riehle, D.M.; Höhenberger, S.; Cording, R.; Delfmann, P. (2017): Live Query - Visualized Process Analysis, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1295-1298

companies require analysis approaches that can be customized to their situation and characteristics. Thus, an objective of a (semi-)automatic model analysis approach should be wide applicability in connection with easy handling. This way, an effective and efficient business process management can be sufficiently supported. This work aims to close this still existing gap.

2 Yet Another Analysis Approach?

Many model analysis approaches make use of model querying. In model querying, an issue to be searched (mostly a compliance violation, an inefficiency or another kind of flaw) is modeled in a specific modeling language and searched within process models by means of a corresponding algorithm. The effectiveness of an approach is based on the assessment if a detected issue structurally and especially content-wise, corresponds to the searched one. The efficiency depicts the required resources (e.g., the needed time). Albeit, their technical efficiency or effectiveness do not predominantly determine their adoption in practice. Their ability to be adapted to a company's needs and their actual usability figure much more prominently. Understandably, companies do (mostly) not employ approaches that do not support the prevalent modeling language, the issues to be searched for in models or that are too difficult to learn for the employees.

Thus, BECKER ET AL. [8] assess different model query approaches regarding their applicability for the detection of a variety of compliance violations. It becomes apparent that only five of the 27 identified approaches can handle different modeling languages. Furthermore, many of the approaches are restricted to search only for issues that exhibit a low level of complexity (i.e., the size and the amount of crosslinking the model subsection to be searched for can have). Some approaches retire at a medium complexity level, and only one approach can represent complex issues. In addition, only eleven approaches have been prototypically implemented, from which only one has been evaluated on real-world process models.

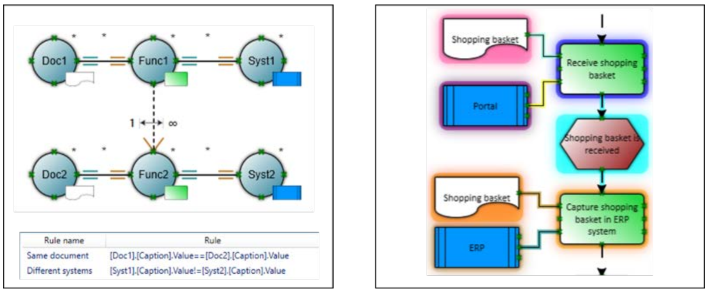
To focus on real-world problems and continuous improvement, we followed the Design Science methodology of PEFFERS ET AL. [9]. We proposed a model query approach under the objectives of wide applicability, independency of modeling languages, support of difficult and complex issues, and, at the same time, an appropriate ease of use [10]. This resulted in the *Generic Model Query Language (GMQL)*, which is based on set theory. Consisting of functions and set operators, it works independently of a modeling language and supports querying complex structural patterns (examples in [8]). An evaluation confirms the power and the utility of *GMQL* but also its amendable ease of use [11]. Though being generally applicable, the approach is inconvenient and overly complex, which prevents the application in practice.

Figure 1. Overall development process

Therefore, starting a new process iteration in our Design Science methodology, we introduced the *Diagramed Model Query Language (DMQL)* to keep the wide applicability of *GMQL* by simultaneously improving the ease of use [12]. With the prototype accompanying this paper, we provide an implementation of our *DMQL* approach. This leads to the overall development process as depicted in Figure 1.

Both approaches, *GMQL* and *DMQL* have been implemented as a plug-in for the meta-modeling tool called [εm]. Similar to *GMQL*, *DMQL* is an intelligent brute-force algorithm, which uses concepts from graph theory (like subgraph isomorphism) to find a query in process models automatically. In contrast, *DMQL* queries are modeled by using edges and vertices that reflect structural relations. Queries depict possibly occurring parts of process models that correspond to the initial real-world issue.

Regarding the main objective, the improvement of the ease of use, *DMQL* enables the visual creation of the queries. Hence, formulating a query in *DMQL* strongly resembles graphical process modeling "as usual." Furthermore, the plug-in supports querying models by using the *DMQL* algorithm with a real-time marking within the process models during the search process. An exemplary model query and a corresponding result are shown in Figure 2. Contextual requirements are included by using captions or attributes of the nodes. A detailed application of the plug-in is shown in the accompanying video. Finally, with the graphical specification of queries, the plug-in implementation is a promising approach to easy-to-use model querying.



Exemplary Query | Returned Result

Figure 2. Exemplary DQML Model Query & Returned Result

3 Conclusion

Summarizing, *DMQL* is an approach to provide a widely applicable and easy to use support for the analysis of (business process) models. It was developed following the design process stated above by passing through several evaluations. A comparison we made between *GMQL* and *DMQL* showed that both are usable independently from a modeling language [13]. In addition, a preliminary evaluation, where a group of modelers was asked to create the same query with both approaches, showed that the average creation time is 20 minutes for *GMQL* and 16 minutes for *DMQL* and modelers have reported to find *DMQL* more convenient, which let us presume a possible ease of use improvement. However, this was only an internal comparison and reliable numbers are still pending. While an in-depth evaluation is still part of future work, *DMQL* addresses the identified research gap with wide applicability and easy handling.

Additionally, the efficiency and effectiveness have been successfully evaluated by [13]. As already done during the past development process, evaluation insights and thereof upcoming ideas are taken continuously into account for further algorithm improvement are an improvement the approach. The latest published version of the approach and the video are accessible via www.conceptual-modeling.org.

References

1. PwC: Zukunftsthema Geschäftsprozessmanagement. (2011).
2. Weske, M.: Business Process Management. Springer, Berlin, Germany (2012).
3. Elgammal, A., Turetken, O., van den Heuvel, W.-J., Papazoglou, M.: Formalizing and Appling Compliance Patterns for Business Process Compliance. Softw. Syst. Model. 1–28 (2014).
4. Sadiq, S., Governatori, G., Namiri, K.: Modeling Control Objectives for Business Process Compliance. In: Proc. of the BPM 2007. pp. 149–164. Brisbane, Australia (2007).
5. Lee, A.S., Baskerville, R.L.: Generalizing Generalizability in Information Systems Research. Inf. Syst. Res. 14, 221–432 (2003).
6. Rosemann, M., Vessey, I.: Toward Improving the Relevance of Inf. Syst. Res. to Practice: The Role of Applicability Checks. MISQ. 31, 1–22 (2008).
7. Höhenberger, S., Riehle, D.M., Delfmann, P.: From Legislation to Potential Compliance Violations in Business Processes - Simplicity Matters. In: Proc. of the ECIS 2016. Istanbul, Türkei (2016).
8. Becker, J., Delfmann, P., Eggert, M., Schwittay, S.: Generalizability and Applicability of Model-Based Business Process Compliance-Checking Approaches - A State-of-the-Art Analysis and Research Roadmap. BuR - Bus. Res. 5, 221–247 (2012).
9. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A Design Science Research Methodology for Inf. Syst. Res. J. Manag. Inf. Syst. 24, 45–77 (2007).
10. Delfmann, P., Steinhorst, M., Dietrich, H.-A., Becker, J.: The Generic Model Query Language GMQL – Conceptual Specification, Implementation, and Runtime Evaluation. Inf. Syst. 47, 129–177 (2015).
11. Becker, J., Delfmann, P., Dietrich, H.-A., Steinhorst, M., Eggert, M.: Business Process Compliance Checking – Applying and Evaluating a Generic Pattern Matching Approach for Conceptual Models in the Financial Sector. Inf. Syst. Front. 1–47 (2014).
12. Delfmann, P., Breuker, D., Matzner, M., Becker, J.: Supporting Information Systems Analysis Through Conceptual Model Query – The Diagramed Model Query Language (DMQL). Commun. Assoc. Inf. Syst. 37, (2015).
13. Delfmann, P., Höhenberger, S.: Supporting Business Process Improvement through Business Process Weakness Pattern Collections. In: Proc. of the Wirtschaftsinformatik 2015. pp. 378–392. Osnabrück, Germany (2015).

Empowering Smarter Fitting Rooms with RFID Data Analytics

Matthias Hauser[1], Matthias Griebel[1], Jannis Hanke[1], and Frédéric Thiesse[1]

[1] University of Würzburg, Department of Business Management, Würzburg, Germany
matthias.hauser@uni-wuerzburg.de, griebelmatthias@gmail.com,
jannis.hanke@uni-wuerzburg.de, frederic.thiesse@uni-wuerzburg.de

Abstract. Smart Fitting Rooms offer great potential to enhance customer shopping experience in retail fashion stores. To this end, they leverage auto-id technology to detect product selections of customers and offer additional services based on these selections on screens within the Smart Fitting Room cabins (e.g., product recommendations or Omnichannel services). While current implementations mainly rely on hardware-based approaches to reliably detect customer's product selection, we investigate the applicability of software-based approaches to distinguish between products in individual cabins. In addition, we show that software-based approaches can be used to sort products in fitting rooms by relevance to the customer which is a valuable information for providing additional services.

Keywords: Data Analytics, RFID, Context Awareness, Internet of Things

1 Smart Fitting Room Opportunities

Internet of Things (IoT) technologies offer great opportunities to enhance customer shopping experience in retail fashion stores [1]. Such context-aware technologies have the potential to create pervasive retail systems that are able to accommodate user needs and wants when desired [2]. We consider Smart Fitting Rooms, which are an example of such context-aware environments. Such fitting rooms are not just cabins for trying on selected garments. Instead, they offer customers additional services on a screen within the cabin based on their product selection. Such services are for example (i) product recommendations or (ii) Omnichannel services.

Product recommendations facilitate cross- and up-selling and can lead to substantial sales increases for retailers [3]. Wong et al. [4], for example, provide such recommendations to customers in Smart Fitting Rooms at a fashion chain store in Hong Kong and find that they lead to sales improvements of more than 20%.

Omnichannel services, on the other hand, provide customers with a seamless shopping experience which can lead to a competitive advantage for retailers [5]. The smart fitting room offers various possibilities to bridge the gap between the different retail channels by, for example, offering customers to purchase products that are currently not available in the store from the online store in the Smart Fitting Room.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Hauser, M.; Griebel, M.; Hanke, J.; Thiesse, F. (2017): Empowering Smarter Fitting Rooms with RFID Data Analytics, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1299-1302

2 Limitations of Current Implementations

Retailers have already started testing fitting rooms that offer additional services to customers. While some of these retailers provide customers with barcode scanners for the identification of items, others rely on Radio Frequency Identification (RFID) technology. RFID uses electromagnetic fields to automatically identify and track tags that are attached to objects. The technology is already commonly used in the retail supply chains for the automatic detection of logistical units in upstream and backroom processes. In contrast to optical barcode scanning, RFID tags not only enable the detection of the amount of items belonging to a specific product category but also permit the identification of each specific item. Moreover, RFID-based object identification does not require a direct line of sight between the tag and the reader device and thus allows for simultaneous bulk detection of multiple objects.

Thiesse et al. [6], Melià-Seguí et al. [7], and Wong et al. [4] describe pilot projects with Smart Fitting Rooms that rely on RFID technology. The application of RFID for such processes is, however, error-prone and challenging [8]. This is because in contrast to controlled processes in other parts of the supply chain, the sheer number and variety of simultaneously moving objects is very high. Complexity is further increased by the manner in which objects are transported, unpredictable customer walking paths, and suboptimal store layouts. To cope with these challenges, existing implementations mainly rely on hardware-based approaches. These approaches require the installation of one RFID reader system for each cabin and the use of shielding measures (e.g., shielding paint, thick fitting room walls from floor to ceiling) to ensure that only objects within the cabins are detected by the RFID systems. Another drawback of current implementations is that they are only able to detect all products within them but cannot distinguish between relevant products (e.g., products that are currently tried on) and others (e.g., products that hang on a coat hook). It is obvious that such information would be valuable for retailers as this would, for example, allow them to display only recommendations for items that customers are currently interested in.

3 Description of the Smart Fitting Room Artifact

We investigate the applicability of software-based approaches to tackle the previously mentioned limitations of current implementations. To this end, we follow the guidelines put forward by Hevner et al. [13] and develop an artifact that (i) detects garments within cabins without using shielding measures and (ii) sorts them by relevance to customers. Our current experimental setup considers three fitting room cabins (Figure 1).

The architecture of the artifact combines hardware and software components. The hardware component is a sensor infrastructure that collects sensor data that is then processed by the software components. We propose the use of a ceiling-mounted RFID system with a reader device and an antenna array with 52 far-field antenna beams which are mounted in one housing. Table 1 provides an exemplary data excerpt from the raw data gathered with the reader installation. Each row reflects a single tag read event triggered by one of the reader's antennas. Here, EPC is the unique identifier of the RFID tag, RSSI is the radio signal's power measured in dBm, Doppler is the frequency shift of the received signal at the reader due to relative motion between the reader and the tag, and Antenna is the ID of the antenna that read the tag.

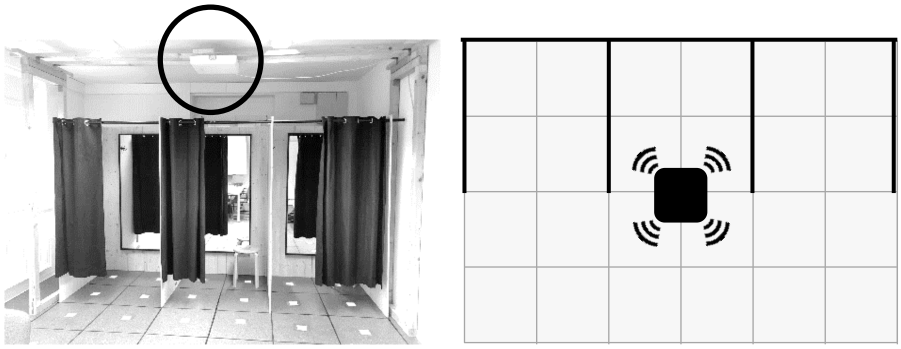


Figure 1. Test environment with ceiling-mounted RFID system and three fitting room cabins

Table 1. Exemplary data excerpt

EPC	***Timestamp***	***RSSI***	***Doppler***	***Frequency***	***Power***	***Antenna***
E28011606…	1465992659	-66.0	25.187	867.5	30.0	51
E28011606…	1465992660	-64.0	30.875	867.5	30.0	2

The first software component combines supervised machine learning and probabilistic models to detect items in individual cabins based on the low-level RFID data collected by the hardware component. In a first step, we determine the positions of tagged items using the indoor positioning technique “Scene Analysis”, which estimates the position of objects by matching their real-time measurements with “fingerprints” at different positions [9]. To this end, we apply multiclass classification techniques which necessitate dividing the fitting room area into grid fields and collecting training data for each of these fields. Here, the number of grid fields denotes the number of classes considered in the data mining model. In a second step, we consider the layout of the fitting room area and characteristics of the processes in fitting rooms to improve the positioning accuracy. The former comprises for example information about the location of fitting room walls, the latter builds on the assumption that some sequences of item locations within a certain time period are more likely than others.

The second software component relies on supervised machine learning techniques to sort garments in fitting rooms (i.e., the garments detected by the first software component) by relevance. To this end, we extract features from the raw data, which contain information regarding observed real-world events. The considered features are appropriately specific to RFID and must be developed based on knowledge of the particular business process. Examples are the number of tag reads, the standard deviation of the RSSI values, or the median of the Doppler frequency shifts of a particular tag within a certain time interval.

4 Expected Contribution and Future Work

Our initial study shows that current limitations of existing implementations can be tackled with software-based approaches. Our artifact is able to detect products in cabins (without the need of shielding measures) and sort these by relevance to the customer. Going forward, we want to evaluate our artifact in different test environments. Here, we want to focus on complex settings, such as high numbers of customers and products in reading range of the antennas or different customer movement speeds. Although first results are promising, we expect that additional tests will introduce new challenges. Our ultimate objective is to ensure feasibility of our approach under real-world conditions.

5 Acknowledgement

This research was carried out under the Research Project SERAMIS (Sensor-Enabled Real-World Awareness for Management Information Systems), funded by the EU under the VII Framework Research Program.

References

1. Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, Jacques, Aharon, D.: The Internet of Things: Mapping the Value Beyond the Hype. McKinsey Global Institute (2015).
2. Kourouthanassis, P., Roussos, G.: Developing Consumer-Friendly Pervasive Retail Systems. IEEE Pervasive Comput. 2, 32–39 (2003).
3. Senecal, S., Nantel, J.: The influence of online product recommendations on consumers' online choices. J. Retail. 80, 159–169 (2004).
4. Wong, W.-K., Leung, S.Y.S., Guo, Z.X., Zeng, X.H., Mok, P.Y.: Intelligent product cross-selling system with radio frequency identification technology for retailing. Int. J. Prod. Econ. 135, 308–319 (2012).
5. Blázquez, M.: Fashion shopping in multichannel retail: The role of technology in enhancing the customer experience. Int. J. Electron. Commer. 18, 97–116 (2014).
6. Thiesse, F., Al-Kassab, J., Fleisch, E.: Understanding the Value of Integrated RFID Systems: A Case Study from Apparel Retail. Eur. J. Inf. Syst. 18, S. 592-614 (2009).
7. Melià-Seguí, J., Pous, R., Carreras, A., Morenza-Cinos, M., Parada, R., Liaghat, Z., De Porrata-Doria, R.: Enhancing the shopping experience through RFID in an actual retail store. In: Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication. pp. 1029–1036. ACM (2013).
8. Hauser, M., Zügner, D., Flath, C.M., Thiesse, F.: Pushing the limits of RFID: Empowering RFID-based Electronic Article Surveillance with Data Analytics Techniques. In: Proceedings of the Thirty Sixth ICIS (2015).
9. Liu, H., Darabi, H., Banerjee, P., Liu, J.: Survey of Wireless Indoor Positioning Techniques and Systems. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. 37, 1067–1080 (2007).

You can't buy my rating! On the pivotal effect of an unconditional gift on rating behavior

Dominik Gutt[1], Darius Schlangenotto[1], Dennis Kundisch[1]

[1] Paderborn University, Faculty of Business Administration and Economics, Paderborn, Germany
{dominik.gutt,darius.schlangenotto,dennis.kundisch}@uni-paderborn.de

Abstract. The importance of online ratings on sales is widely acknowledged. Firms need to find ways of increasing the number of ratings and rating scores, but how they can achieve this effectively is less well established. In this paper we analyze the impact of an unconditional gift on customers' rating behavior in an online field experiment. Contrary to prevalent advice, our results suggest that providing a gift is not necessarily beneficial. Younger customers are significantly less likely to rate when exposed to an unconditional gift. Regression analysis reveals that age serves as a moderator and older customers even respond slightly positive to a gift. Having detected a negative effect of gifts on rating behavior provides first indicative evidence of a possible crowding out of intrinsic motivation in the context of online ratings. This has direct implications for practitioners considering the usage of gifts to elicit online ratings.

Keywords: online ratings, rating elicitation, reciprocal behavior, field experiment

1 Introduction

How many times have you received an email asking you to rate your recent online purchase but not acted on it? If your answer is "often", then you are in good company. Such email solicitations may be increasingly common but review rates remain stubbornly low, typically within one-digit percentage success rates (1.5%, as reported by Anderson and Simester [1]). Given that customer online ratings are a major driver of purchase behavior both online [2] and offline [3], by making it easier for customers to evaluate and compare the quality of a product or service, it is in the interest of businesses to increase review rates. This is supported by a considerable body of literature which, by and large, lends empirical evidence to the positive impact that online ratings have on sales in a variety of industries. This positive relationship is primarily driven by the volume of ratings [4] and the average ratings [5]. Consequently, obtaining a substantial number of ratings and achieving high average ratings has become a critical endeavor for firms both in online and offline markets.

Studies in the offline world have provided empirical evidence in support of the claim that customer feedback can be successfully elicited through monetary and nonmonetary

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Gutt, D.; Schlangenotto, D.; Kundisch, D. (2017): You can't buy my rating! On the pivotal effect of an unconditional gift on rating behavior, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1303-1317

gifts, increasing both the amount of feedback [6], [7] as well as the average feedback value [8], [9]. Unconditional gifts in the offline world can be cash payments [6], [7]; [10], lottery tickets [7], charitable donations on behalf of the respondent [7], or fancy sweets [8]. While there is substantial evidence that gifts can increase and enhance feedback in offline environments, only a few recent studies involving field experiments on eBay have investigated how sellers can effectively elicit ratings in the e-commerce environment [11], [12]. Up to now it remains unclear, however, whether the results produced in the context of auction markets are transferable to conventional markets. In addition, the researchers in the aforementioned studies made use of *conditional* gifts which customers receive only in exchange for submitting a rating. For example, selling USB sticks on eBay, they employ pre-announced price discounts before purchase *conditional* on the customer’s rating [12]. Moreover, a related study found that the effect of conditional rebates on rating behavior and sales varies with the amount of discounts [11]. However, research findings from the offline world suggest that *unconditional* gifts offered to customers regardless of their subsequent action might be even better suited to elicit ratings [10].

This study, then, aims to analyze the impact of unconditional gifts on the rating behavior of customers in a conventional e-commerce environment. Unconditional rebates are provided as a gift via post-purchase emails in which customers are asked to rate their purchase. Investigating ways to gather and maintain good customer ratings is crucial for practitioners and scholars alike. This is especially true of post-purchase emails, which represent a cost-efficient - and now widespread - tool to obtain ratings. Thus we pose the following research question: *How do unconditional gifts in email elicitations affect the online rating behavior of customers?*

Previous insights on eBay [12] focused exclusively on the effectiveness of *conditional* rebates on rating behavior. With our research question we attempt narrowing the knowledge gap concerning *unconditional* rebates in a different market setting. Therefore we conduct an online field experiment and find a substantially heterogeneous treatment effect in response to unconditional gifts. Evidence suggests that emails offering unconditional gifts significantly decrease the rating volume provided by customers. This effect is moderated by customer age. While the first and second quartiles of the age distribution give significantly fewer ratings when receiving a treatment, the effect gradually lessens for the third and fourth age quartiles. Moreover, this effect is more pronounced for recurring than for new customers. Finally, we find no empirical evidence for unconditional gifts influencing the average rating.

Thus our research makes several contributions to the literature and carries valuable implications for scholars and practitioners alike. First, we add to the literature on online rating elicitation by broadening the scope from an online auction environment to conventional online business-to-consumer (B2C) commerce. Second, we add to the literature on reciprocal gifts by presenting empirical evidence suggesting that customers might perceive a gift as an attempt by a firm to influence their rating behavior, against which they then react in the form of a decrease in the rating volume. Finally, our results enable us to derive practical managerial implications. Managers intent on increasing rating volumes should be aware that email elicitations offering unconditional gifts do not automatically result in an increase in the number of customers providing ratings.

Rather, in the younger customer base (in our case, aged between 18 and 48) the emailed gift offer has the effect of decreasing the rating volume, whilst only slightly increasing the response rate of the older customers (aged 49-85). Managers might therefore want to design marketing interventions capable of exploiting this observed behavior by targeting them specifically at older customers.

2 Related Literature

A sizeable and emerging body of literature provides empirical evidence for the relationship between online ratings and business performance. One sub-stream identified the rating volume as one key determinant. The rating volume has been found to be positively associated with sales for books on Amazon [4] and movies on the basis of a variety of data sources [13]. Another sub-stream identified the rating valence as a crucial determinant. The rating valence, measured as the average rating obtained through ratings, has been found to be positively associated with revenues of a restaurant [3] and with sales in the online book market [5], [14] and the movie industry [15].

In the offline world a sizeable body of literature provides empirical evidence that customer response rates can be increased by offering monetary as well as nonmonetary gifts. Gifts (such as cash payments, lottery tickets, charitable donations, or sweets) have been found to increase the feedback volume [6], [7] as well as the feedback value [8], [9]. Singer et al. [6] find a positive effect of unconditional prepaid cash gifts on mail survey response rates. Warriner et al. [7] state that unconditional prepaid cash gifts can be even more effective than alternative forms as gifts such as lottery tickets or charitable donations. Moreover, Strohmetz et al. [8] find that unconditional chocolate gifts can significantly increase the tips given by customers. Comparing monetary gifts, James and Bolstein [10] conclude that an unconditional gift of $5 is more effective at increasing the volume of responses to a mail survey than a $50 conditional gift. In the online world, researchers have only started to investigate ways in which online ratings can be actively elicited. The current body of literature comprises research studies conducted in auction environments. Cabral and Li [12] argue that conditional gifts can trigger the reciprocal behavior of buyers and are thereby result in an increased rating volume and an enhanced average rating. In a related study, Li and Xiao [11], for example, conclude that conditional rebates lead to an increase in the number of sales and in the likelihood of obtaining a good rating.

We contribute to this literature by extending the investigation of rating elicitation towards unconditional gifts and towards a more generalizable non-auction e-commerce context. The impact of unconditional gifts on customer behavior is measured in terms of rating volume and average rating.

3 Theoretical Background and Hypotheses

Before analyzing the effect of an unconditional gift on rating behavior, we will review the literature on rating behavior in general, and specifically the role that gifts can play to influence such behavior.

One widely accepted stream of literature recognizes the positive relationship between the net utility a customer derives from the consumption of a product and the volume and average ratings they provide [2], [16], [17]. Theoretically, customers tend to publish a rating equal to the utility they derive from the obtained good. Yet, they will only provide a rating if they perceive its utility to be greater or equal to zero considering the costs incurred by the rating activity, such as the time needed to reflect on and compose a review [12]. However, observations in the field suggest that far from all customers who derive a positive utility give an online rating, and that ratings are given when the utility is either very high or very low [18].

The theory of reciprocity suggests that humans intuitively feel obliged to give back (reciprocate) in response to the actions of others. When a gift is presented to customers, reciprocal theory suggests that they will evaluate the gift and alter their behavior based on the psychological utility they derive from it. According to Falk and Fischbacher [19], utility from a reciprocal gift consists of three components. It comprises (i) the customer's perception of the gift-giver's intention, (ii) the value of the gift minus the gift-giver's outside option, and (iii) the reciprocation. Intention captures the notion that gifts can be based on disingenuous intentions, which hence lowers the value of the gift. The value of the gift minus the gift-giver's outside option reflects the differential between the value of the gift minus what the giver could have given. Reciprocation captures the value of the gift that prompts a reciprocally-acting individual's action by them returning the favor or rejecting it. Thus, reciprocity works both ways: customers respond positively to gifts perceived as genuine incentives and negatively to gifts that are seen as manipulative or disingenuous.

The total perceived utility a customer derives is constituted by the physical utility of consumption (the purchase and use of the contact lenses) plus the psychological utility due to the reciprocal gift which, in turn, consists of three components. From the perspective of a customer, one perception of a reciprocal gift could be that the customer (i) perceives the gift as genuine, (ii) values the gift and considers the gift giver's outside option as zero (because the retailer could have as well opted to give nothing at all), and (iii) reciprocates the kind behavior by giving a rating. This would result in a positive psychological utility and positive reciprocal behavior. The part of the utility derived by (iii) reciprocation is based on the idea that by giving something back, you feel better because you act fair [19] and thus you derive utility from acting reciprocally after receiving a gift. In that case, the perceived utility in the presence of the reciprocal gift is higher than in the absence of the gift.

In sum, current literature suggests that both rating volume and average ratings are driven by the perceived utility a customer derives from the obtained good. Introducing an unconditional gift to elicit ratings could either increase or decrease this perceived utility. In our research environment the gift should increase the overall utility in the minds of customers as they receive the gift in form of a rebate which was designed to be attractive in terms of its monetary value and which does not depend on a customer's action. Thus, we formulate our first set of hypotheses:

H1a: An unconditional gift increases the number of ratings received by the seller.

H2a: An unconditional gift increases the average rating received by the seller.

However, as stated by Falk and Fischbacher [19], customers could also perceive a gift as negative and respond to it accordingly, i.e. by refusing to act as requested. In our research environment, even when the gift does not depend on a customer action and is also considered attractive in terms of monetary value, customers might interpret the intentions of the gift-giver as being disingenuous. They might get the impression that the gift-giver wants to buy their rating. In these cases, the perceived utility of the gift could become negative, resulting in a negative impact of the gift on rating behavior. Consequently, we formulate a competing set of hypotheses:

H1b: An unconditional gift decreases the number of ratings received by the seller.

H2b: An unconditional gift decreases the average rating received by the seller.

4 Research Design

To evaluate the impact of an unconditional gift on rating behavior we teamed up with a German B2C contact lens retailer who uses emails to elicit online ratings from customers. Every customer who bought items in their shop receives a rating elicitation email three days after their items have been shipped. Obtained ratings are publicly accessible and can be viewed by all potential shoppers. We designed a well-controlled field experiment which leveraged the firm's practice of sending out rating elicitation emails, allowing us to identify the causal effect of an unconditional gift on rating behavior. As can be seen in Figure 1, whenever customers purchase items in the online contact lens shop they are randomly assigned to either the control or the treatment group. Customers assigned to the control group receive an email that asks them to rate their customer experience. When assigned to the treatment group, customers receive a modified email which presents an unconditional rebate as a gift. In order to ensure proper randomization, the assignment to either of the two conditions (treatment or control) is performed automatically. In every elicitation email, customers have the distinct choice to either rate the firm via a third-party website or refrain from rating (tested via hypothesis H1a/H1b). Customers who decide to give a rating can assign between one to five stars to the distinct categories, namely, delivery, product and service (tested via hypotheses H2a/H2b).

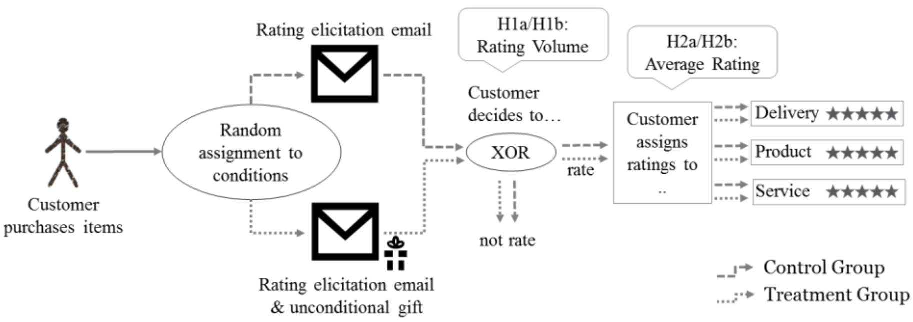


Figure 1. Research Setup

During the experimental period customers were equally likely to either receive the regular elicitation email (see Figure 2, Control) or a modified elicitation email which additionally offers an unconditional gift (see Figure 2, Treatment). While the subject ("Rate 'brand name' now!") of both emails was kept identical to rule out potential biases, the email content was altered in two ways. First, the headline differed: the treatment email only mentions the gift offered to the customer (see Figure 2, Treatment – [A]), whereas the control email only asks customers about the service they received ("how was our service?"). Second, below the rating-button an additional descriptive text was inserted in the treatment email, stating the importance of customer satisfaction to the firm and introducing the gift offer of five Euros in form of a rebate for the next purchase (see Figure 2, Treatment – [B]). These two modifications are used to advertise the unconditional gift and to trigger the reciprocal behavior pattern in customers. We chose a rebate of five Euros as it is considered to represent a substantial rebate on the average purchase basket. In addition, previous marketing campaigns have proven that customers of the contact lens shop respond positively to rebates and actively use them. Yet, in our research context no previous campaigns had tried to incentivize ratings via gifts. This means that prior to our intervention, the firm's customers would not have expected to receive a rebate in a rating elicitation email.

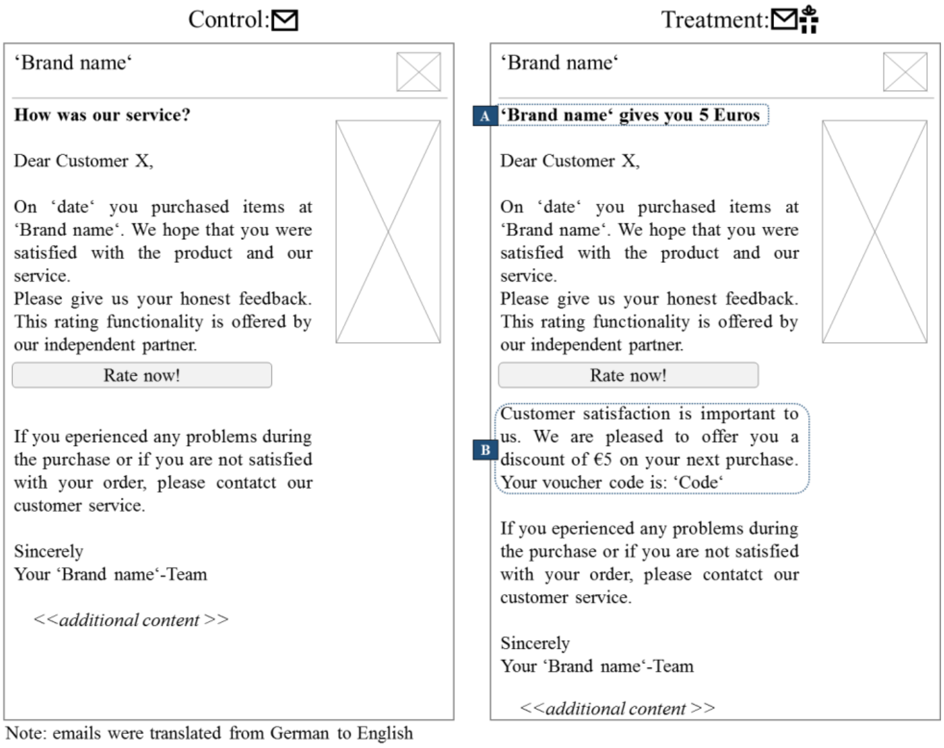


Figure 2. Rating Elicitation Emails

5 Empirical Model

Regression analysis is used to investigate the gift effect while simultaneously controlling for other potential confounding factors. Table 1 lists the variables considered in our regression analysis. Each customer *i* is randomly assigned to either a control group or a treatment group ($TREATMENT_i$) receiving a differently worded elicitation email (see Figure 2). We subdivide the sample of customers by their age so that every customer *i* is assigned to one of four 25%-quartiles (AGE_i) of the age distribution (see Figure 3). Customers are further distinguished by gender ($GENDER_i$) and by customer status ($CUSTOMER_i$). A customer could either be a new or a recurring customer. The rating behavior of every customer *i* is assessed first by whether they decided to provide a rating ($RATED_i$), and second, for those who did, whether they rated the category 'delivery' ($DELIVERY_i$), 'product' ($PRODUCT_i$) or customer 'service' ($SERVICE_i$) as the dependent variables of interest.

Table 1. Main Variables

Variable	Type	Description	Value range
$TREATMENT_i$	Independent	Every customer *i* is assigned to the control or treatment group	CONTROL/ TREATMENT
AGE_i	Control	Based on his age every customer *i* is assigned to an age quartile	Q1/Q2/Q3/Q4
$GENDER_i$	Control	Every customer *i* is either male or female	MALE/FEMALE
$CUSTOMER_i$	Control	Every customer *i* is either classified as a new or a recurring customer	RECURRING/ NEW
$RATED_i$	Dependent	Every customer *i* either rates (=1) or refuses to rate (=0)	[0,1]
$DELIVERY_i$	Dependent	Every customer *i* who rates can assign a rating to the quality dimension *delivery*	[1,5]
$PRODUCT_i$	Dependent	Every customer *i* who rates can assign a rating to the quality dimension *product*	[1,5]
$SERVICE_i$	Dependent	Every customer *i* who rates can assign a rating to the quality dimension (customer) *service*	[1,5]

By including all customer attributes as controls, we can distinguish between effects that are caused by the treatment and those that are driven by age, gender and customer type. Furthermore, previous studies suggest that a customer's age might be a pivotal determinant of their reaction to external marketing stimuli [20, 21]. We consequently include the interaction term between AGE_i and $TREATMENT_i$ to test for age-specific group differences. Hence we formulate the following model which is used as a logit regression for the dichotomously distributed rating volume and as an OLS regression to assess the average ratings:

$$Y_i=\beta_0+\beta_1 TREATMENT_i+\sum_{\tau=1}^{3}\beta_\tau\ AGE_i+\beta_3 GENDER_i+\beta_4 CUSTOMER_i+\sum_{\tau=1}^{3}\beta_{\tau i}\ (AGE_i{*}TREATMENT_i)+\varepsilon_i \quad (1)$$

6 Empirical Analysis

Throughout the 67 days of the experimental period, a total of 7,316 customers received elicitation emails three days after their products were shipped. As can be seen in Table 2, the majority of customers were recurring customers (78% or 5,697). Based on observable attributes, every customer was randomly assigned to either the control or the treatment group. To test whether the randomization of the treatment assignment worked properly, we performed a Mann-Whitney-Wilcoxon-test. The test yielded no significant group differences based on the attributes: customer type (recurring / new) and gender (male / female). Therefore, we find no indication to doubt the randomization applied in our experiment. Figure 3 depicts the age distribution of all customers, ranging from 18 to 85 years.[1] For every quartile an age group is built to allow for the observation of age-specific treatment effects.

Table 2. Received Elicitation Emails

Group/ Customer Attributes	Control	Treatment	Total
Recurring Customer	**2,820**	**2,877**	**5,697**
female	1,912	1,907	3,819
male	908	970	1,878
New Customer	**824**	**795**	**1,619**
female	558	569	1,127
male	266	226	492
Total	3,644	3,672	7,316

Table 3 presents customers' rating behavior during the experiment in respect of age quartiles. During the experiment 5.2% (379) of all customers chose to provide a rating. Their likelihood to rate increases with AGE, from 2% for the youngest quartile to 9.6%

[1] 55 customers had to be removed from the dataset as their age could not be determined.

for the oldest. These rating differences underline the need to differentiate customer behavior by age. In regard to the average rating given, customers tend to assign ratings close to the maximum value of five.

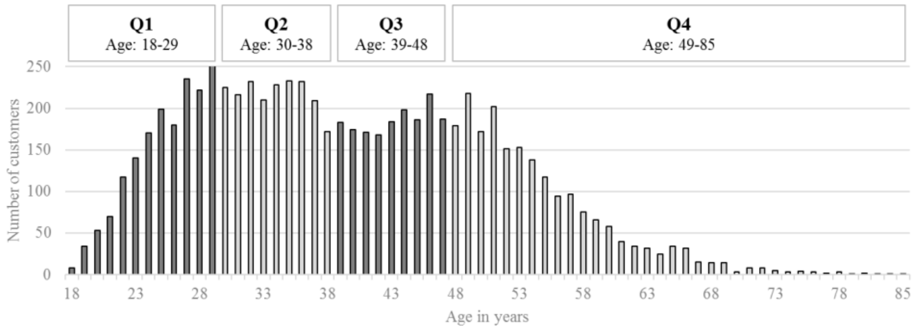


Figure 3. Age Distribution

For all the rating categories, ratings range from 4.7 to 4.9 (this is in line with high overall ratings as reported by [18]). In order to assess whether this observed rating behavior is influenced by the unconditional gift, the rating volume and the average ratings are analyzed separately.[2]

Table 3. Rating Behavior

Behavior / Age Quartiles	Rating Volume	Average Rating		
		Delivery	Service	Product
Q1	2.0%	4.9	4.8	4.8
Q2	3.1%	4.7	4.7	4.7
Q3	5.6%	4.9	4.9	4.9
Q4	9.6%	4.8	4.8	4.9
Total	5.2%	4.8	4.8	4.9

Notes: Rating Volume is calculated as: Number of ratings / Received elicitation emails
Average Rating is calculated for every dimension as: Number of assigned rating stars/Number of ratings

6.1 Rating Volume

Table 4 depicts customers' rating volume by their individual attributes. On average customers from the control group rated the firm in 5.5% of all instances as opposed to customers in the treatment group who provided a rating in 4.9% of all cases, indicating that the overall treatment effect is negative. The negative treatment effect is observed

2 Additional statistical analyses, extended variants of all regression models, incl. control variables are provided as an online resource: http://go.upb.de/Reciprocity

for recurring as well as for new customers. Table 5 presents the results of logit regressions performed according to the regression model specified in Equation (1).[3]

As can be seen in column (1) of Table 5, the treatment effect on the first quartile, comprising customers aged between 18 and 29, is significantly negative with a coefficient of -0.80. As logistic regression analyses are applied, coefficients cannot be interpreted as the direct impact on a change in the output variable for a one-unit increase in the respective predictor variable, while all other predictors remain constant. Instead, odds-ratios are more appropriate [22]. The odds-ratio indicates that the youngest group of customers in the data set are 55% less likely to provide a rating when exposed to an unconditional gift. In column (2), customers between the ages of 30 and 38 are also significantly less likely to respond to the gift with a rating, given their quartile's coefficient of -0.52. The odds-ratio here suggests that the second youngest group of customers is 41% less likely to rate in response to the gift. Customers between the ages of 39 and 48 also seem to be affected negatively, but the treatment effect is insignificant for this quartile. Only the oldest quartile of customers is not influenced negatively by the gift with logit results indicating a rather positive – if ever so slightly positive - response to the treatment.

Table 4. Rating Volume – Summary Statistics

Group/ Customer Attributes	Control (C)	Treatment (T)	Group Differences T-G
Recurring Customer	**5.2%**	**4.9%**	**-0.3%**
female	5.4%	4.5%	-0.9%
male	4.9%	5.5%	0.6%
New Customer	**6.7%**	**5.0%**	**-1.7%**
female	6.1%	4.7%	-1.4%
male	7.9%	5.7%	-2.2%
Total	5.5%	4.9%	-0.6%

From this we can conclude that customers' decision to provide a rating is indeed affected by the emailed offer of an unconditional gift, but in unexpected and differentiated ways: for the majority of customers, the gift not only fails to act as an incentive but rather as a deterrent, since these customer are less likely to submit a rating. The odds-ratio for customers who belong to the younger half of the sample suggests that they are 46% less likely to rate when they receive a reciprocal gift, compared with the control group. Consequently, based on the coefficients of TREATMENT in Table 5, we reject hypothesis H1a and accept the competing hypothesis H1b. Moreover, the interaction terms reveal a substantial heterogeneity of the treatment effect with respect

[3] To evaluate interaction effects between every age quartile and the treatment, four logit regressions are reported in which the particular quartile of interest is used as the respective base case. For example, in the column "Age Q1" the first age quartile is the base case.

to age. For age quartiles Q1, Q2 and Q3 the interaction effect remains insignificant which suggests that there are no significant differences in the treatment effect among the first three age quartiles. However, when comparing the first three quartiles to the fourth (see Table 5, TREATMENT*Q4) logit results yield significant group differences for Q1 and Q2. In column (1) of Table 5, comparing the treatment effect on the youngest quartile to the oldest, the latter significantly differs with a coefficient of 0.95 relative to the youngest quartile. In other words, the customers in the oldest quartile of the sample respond significantly differently. This effect is consistent across columns (1) and (2). In sum, our analysis of the interaction effects reveals that age is a pivotal determinant of a customer's decision to provide a rating in response to receiving an unconditional gift offer by email.

Table 5. Rating Volume – Regression Results

	Dependent Variable: Rated			
Model	Age Q1 (1)	Age Q2 (2)	Age Q3 (3)	Age Q4 (4)
TREATMENT	**-0.80****	**-0.52***	**-0.26**	**0.15**
	(0.38)	(0.27)	(0.21)	(0.15)
TREATMENT*Q1		-0.28	-0.54	-0.95**
		(0.47)	(0.44)	(0.41)
TREATMENT*Q2	0.28		-0.27	-0.67**
	(0.47)		(0.34)	(0.31)
TREATMENT*Q3	0.54	0.27		-0.40
	(0.44)	(0.34)		(0.26)
TREATMENT*Q4	0.95**	0.67**	0.40	
	(0.41)	(0.31)	(0.26)	
Constant	-3.68***	-3.27***	-2.77***	-2.39***
	(0.22)	(0.17)	(0.15)	(0.12)
Controls	✓	✓	✓	✓
Observations	7,316	7,316	7,316	7,316
Log Likelihood	-1,419	-1,419	-1,419	-1,419
Akaike Inf. Crit.	2,858	2,858	2,858	2,858
Notes:	*p<0.1; **p<0.05; ***p<0.01; Standard errors in parentheses; Controls: AGE, GENDER, CUSTOMER			

Average Rating We performed a series of robustness checks. The results remained qualitatively unchanged when (i) we used different age groupings such as a median split, a fixed 20-year and a fixed 10-year interval, and (ii) we conducted estimations separately for recurring customers only. When restricting the logit regression to recurring customers, the observed negative treatment effect increases even further. In other words, recurring customers belonging to the younger half of the sample responded even more negatively to the treatment.[4]

6.2 Average Rating

Table 6 lists the average ratings appertaining to the quality dimensions of ratings provided on delivery, service and product. Group differences are close to zero and range from -0.3 for recurring male customers to + 0.3 for new male customers. Regression results do not yield any significant treatment effect either, as no differences in behavior based on a customer's age can be detected. In our experiment, an unconditional gift only affects the rating volume but has no effect at all on average ratings. Consequently, hypotheses H2a and H2b are rejected as no effect of TREATMENT on ratings can be found for any of the four age quartiles.

Table 6. Average Ratings – Summary Statistics

Group/ Customer Attributes	Control (C)			Treatment (T)			Group Differences
	Delivery	Service	Product	Delivery	Service	Product	∑T- ∑C
Recurring Customer	**4.8**	**4.8**	**4.8**	**4.8**	**4.7**	**4.8**	**-0.1**
female	4.9	4.8	4.9	4.9	4.8	4.9	0.0
male	4.8	4.8	4.7	4.7	4.6	4.7	-0.3
New Customer	**4.8**	**4.7**	**4.9**	**4.8**	**4.9**	**4.9**	**0.2**
female	4.8	4.8	4.8	4.7	4.9	4.8	0.2
male	4.8	4.6	4.9	4.8	4.8	5.0	0.3
Total	4.8	4.8	4.9	4.8	4.8	4.9	-0.1

7 Discussion

There is strong empirical evidence in the literature that the volume and the average of online ratings exert a positive effect on sales. Yet, little is known about how to elicit online ratings from customers effectively and the literature has only just begun to investigate this topic. This paper attempts to fill this gap in two ways: First, we add to the literature on rating elicitation by presenting results from conventional e-commerce markets, thus enhancing prior work that focused mainly on auction markets [11], [12]. Second, we shed light on the effectiveness of rating elicitations via unconditional gifts by means of our empirical field experiment. In general, our rating behavior analysis indicates that especially older customers (in our case, those over 49) are more likely to

[4] Regression results for all robustness checks are provided in our online appendix: http://go.upb.de/Reciprocity#robustnessChecks

rate firms and respond positively to an email elicitation. Rating volumes nearly quadruple between the youngest and oldest quartile of customers. However, our results suggest that providing unconditional gifts via emails fails to elicit ratings. On the contrary, such unconditional gifts tend to decrease the number of ratings, on average, and do not affect the ratings given. The observed negative impact on rating volume is substantially heterogeneous across age quartiles. Regression analyses reveal that the negative effect is mainly driven by the younger half of all customers who become 46% less likely to rate when they receive a gift in the form of an unconditional rebate offer. One potential explanation might be that, at least in our context, gifts (as external stimuli) crowd out the rater's intrinsic motivation. Crowding out of intrinsic motivation implies that utility can be constituted not only by consumption and by reciprocity, but additionally also by other individual-specific aspects or motives (e.g., helping potential customers, expressing themselves in public, showing power over the producer [23]). Unconditional gifts in the form of money might potentially erode these motives leading people to abstain from rating who would have otherwise rated due to their intrinsic motivation. The crowding out effect of intrinsic motivation has received quite substantive scholarly attention [24], [25]. Additionally, effort to redeem the voucher or the value of the voucher itself could contribute to the differences in rating behavior across age groups. However, to the best of our knowledge, this effect has not been observed in the context of online rating elicitation. In our field experiment, intrinsic motivation could be crowded out as raters might feel bribed by the gift and therefore refuse to provide a rating. From the perspective of reciprocity theory introduced in chapter 3, the negative treatment effect might be explained by the fact that costumers sense disingenuous intentions with the retailer for giving out rebates. Thus, the customers retaliate, i.e., abstain from rating. However, it is important to note that we have given out the gift unconditionally, which is not as "aggressive" as a conditional gift and should therefore mitigate the feeling of being bribed.

The results presented in this paper offer straightforward managerial implications and the potential impact can be substantial. Given the fact that spending on e-commerce advertisement has more than tripled during the past five years [26], e-commerce retailers need to find ways to economize on these expenses. Nevertheless, rebates and coupons sent out via emails are currently the preferred way of gift-giving to customers [27] and more than 78% of customers use rebate coupons more than once a year [28]. Therefore, discovering efficient ways to conduct rating elicitation via email transaction with customers can have a crucial impact for e-commerce retailers. Our evidence of a negative effect of unconditional gifts on rating behavior suggests that firms should resist the temptation of using gifts as a way of eliciting ratings. Our results point towards a possible crowding out of intrinsic motivation, with gifts having the exact opposite effect on the intended rating behavior. However, as our experimental data does not allow us to determine the root causes of this customer behavior, future research is needed to establish customers' true intentions. In general, though, firms considering the use of external marketing stimuli should be cognizant of our finding that customer age could be a pivotal determinant in behavioral responses to an external stimulus. At least in our case, customer age plays a decisive role and experimental evidence suggests that older customers react slightly more positively to gifts compared to younger customers.

Furthermore, as our results show that gifts significantly influence rating behavior, in one way or another, future research could evaluate alternative gifts such as nonmonetary incentives as a more effective instrument for eliciting rating behavior. As any research, this work also comes with limitations. Contact lenses are highly standardized and frequently bought by a single person, as opposed to a PC or a digital camera. Thus, our results are potentially limited to less complex repeat-purchase goods. Additionally, future research could also vary the shipping time of the product and the delivery of the elicitation mail to investigate possible effects of this relationship.

8 Acknowledgements

This work was partially supported by the German Research Foundation (DFG) within the Collaborative Research Centre “On-The-Fly Computing” (SFB 901).

References

1. Anderson, E.T., Simester, D.I.: Reviews Without a Purchase. Low Ratings, Loyal Customers, and Deception. Journal of Marketing Research 51, 249–269 (2014)
2. Cabral, L., Hortacsu, A.: The Dynamics of Seller Reputation. Evidence from ebay. The Journal of Industrial Economics 58, 54–78 (2010)
3. Luca, M.: Reviews, Reputation, and Revenue: The Case of Yelp.com. Harvard Business School Working Paper 12-016 (2011)
4. Chevalier, J.A., Mayzlin, D.: The Effect of Word of Mouth on Sales. Online Book Reviews. Journal of Marketing Research 43, 345–354 (2006)
5. Li, X., Hitt, L.M.: Self-Selection and Information Role of Online Product Reviews. Information Systems Research 19, 456–474 (2008)
6. Singer, E., Hoewyk, J., Maher, M.P.: Experiments with Incentives in Telephone Surveys. Public Opinion Quarterly 64, 171–188 (2000)
7. Warriner, K., Goyder, J., Gjertsen, H., Hohner, P., McSpurren, K.: Charities, No; Lotteries, No; Cash, Yes. Main Effects and Interactions in a Canadian Incentives Experiment. Public Opinion Quarterly 60, 542–562 (1996)
8. Strohmetz, D.B., Rind, B., Fisher, R., Lynn, M.: Sweetening the Till. The Use of Candy to Increase Restaurant Tipping. Journal of Applied Social Psychology 32, 300–309 (2002)
9. Gruner, S.: Reward good customers. In: Inc. Magazine, p. 84 (1996)
10. James, J.M., Bolstein, R.: Large Monetary Incentives and Their Effect on Mail Survey Response Rates. Public Opinion Quarterly 56, 442 (1992)
11. Li, L., Xiao, E.: Money Talks. Rebate Mechanisms in Reputation System Design. Management Science 60, 2054–2072 (2014)
12. Cabral, L., Li, L.: A Dollar for Your Thoughts. Feedback-Conditional Rebates on eBay. Management Science 61, 2052–2063 (2015)
13. Duan, W., Gu, B., Whinston, A.B.: Do online reviews matter? An empirical investigation of panel data. Decision Support Systems 45, 1007–1016 (2008)
14. Sun, M.: How Does the Variance of Product Ratings Matter? Management Science 58, 696–707 (2012)

15. Chintagunta, P.K., Gopinath, S., Venkataraman, S.: The Effects of Online User Reviews on Movie Box Office Performance. Accounting for Sequential Rollout and Aggregation Across Local Markets. Marketing Science 29, 944–957 (2010)
16. Herrmann, P., Kundisch, D., Zimmermann, S., Nault, B.: How do Different Sources of the Variance of Consumer Ratings Matter? In: Proceedings of the International Conference on Information Systems (ICIS), Fort Worth (2015)
17. Kwark, Y., Chen, J., Raghunathan, S.: Online Product Reviews. Implications for Retailers and Competing Manufacturers. Information Systems Research 25, 93–110 (2014)
18. Hu, N., Pavlou, P.A., Zhang, J.: Why do Online Product Reviews have a J-shaped Distribution? Overcoming Biases in Online Word-of-Mouth Communication. Communications of the ACM 52, 144–147 (2009)
19. Falk, A., Fischbacher, U.: A theory of reciprocity. Games and Economic Behavior 54, 293–315 (2006)
20. Phillips, L.W., Sternthal, B.: Age Differences in Information Processing. A Perspective on the Aged Consumer. Journal of Marketing Research 14, 444–457 (1977)
21. Zaltman, G., Srivastava, R.K., Deshpande, R.: Perceptions of Unfair Marketing Practices. Consumerism Implications. Advances in Consumer Research 5, 247–253 (1978)
22. Hosmer, D.W., Lemeshow, S., Sturdivant, R.X.: Applied logistic regression. Wiley, Hoboken, New Jersey (2013)
23. Hennig-Thurau, T., Gwinner, K.P., Walsh, G., Gremler, D.D.: Electronic word-of-mouth via consumer-opinion platforms. What motivates consumers to articulate themselves on the Internet? Journal of Interactive Marketing 18, 38–52 (2004)
24. Benabou, R., Tirole, J.: Intrinsic and Extrinsic Motivation. Review of Economic Studies 70, 489–520 (2003)
25. Bowles, S., Polanía-Reyes, S.: Economic Incentives and Social Preferences. Substitutes or Complements? Journal of Economic Literature 50, 368–425 (2012)
26. Axel Springer: Werbeausgaben im E-Commerce in Deutschland in den Jahren 2000 bis 2015 (in Millionen Euro)., https://de.statista.com/statistik/daten/studie/196993/umfrage/werbeausgaben-im-e-commerce-in-deutschland-seit-2000/ (Accessed: 25.10.2016)
27. AffiliPRINT: Welche Gutschein-Typen, egal ob Sie diese erhalten oder nicht, sind für Sie besonders nützlich?, https://de.statista.com/statistik/daten/studie/252314/umfrage/umfrage-zum-interesse-an-ausgewaehlten-gutschein-typen/ (Accessed: 25.10.2016)
28. Statista: Wie oft nutzen Sie Rabatt-Coupons bzw. -gutscheine (z.B. 10% Preisnachlass, Barrabatt, kostenloser Lieferservice)?, https://de.statista.com/statistik/daten/studie/601687/umfrage/haeufigkeit-der-nutzung-von-rabatt-coupons-gutscheinen-in-deutschland/ (Accessed: 25.10.2016)

Setting the Right Tone: How Data Science Enables Investor Communication to Choose the Right Language

Simon Alfano[1], Nicolas Pröllochs[1], Joscha Märkle-Huss[1], Stefan Feuerriegel[1], and Dirk Neumann[1]

[1] University of Freiburg, Chair for Information Systems Research, Freiburg, Germany
{simon.alfano,nicolas.proellochs, joscha.maerkle-huss,stefan.feuerriegel,dirk.neumann}@is.uni-freiburg.de

Abstract. Facts matter in financial communication, but academic research has revealed that soft information such as the tone and readability also impact the decision-making of investors and ultimately stock returns. Information Systems (IS) research has developed measures to quantify readability and tone of textual content, among which are techniques from statistical learning which are capable of extracting the most relevant words for investors. To make this knowledge accessible, we develop an IS prototype that guides communication practitioners to set the right tone in their investor communication. First, we identify relevant practitioner needs using requirements engineering. Second, we translate these needs into an IS prototype which seamlessly integrates into text editing processes and serves as a decision support system to steer readability and tone. Third, we pilot our IS prototype with 37 companies and successfully validate our prototype's capabilities in a survey with financial professionals.

Keywords: Decision Support Systems (DSS), Information Processing, Investor Communication, IS Prototype, Sentiment Analysis

1 Introduction

Financial markets are facing ever increasing information flows [1], [1]. On the one hand, this burgeoning data volume enables analysts to make more educated judgments on valuation and, thus, allows investors to make better informed investment decisions. On the other hand, the sheer amount of available information can be overwhelming for both individuals and companies [1], [1, 3]. Consequently, this inevitably calls for decision support systems in financial markets. Such decision support systems should be directed at all stakeholders involved in the dissemination of information in financial markets. By that, these systems should assist both senders of information (e.g. investor relations departments within companies, financial news agencies) and recipients of relevant information, such as financial analysts and investors [4], [5]. In fact, the latter group, i.e. investors as recipients of relevant information, already relies to a large extent on decision support systems. Examples include forecasting systems, text mining software and trading algorithms which support or improve decision-making.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Alfano, S.; Pröllochs, N.; Märkle-Huss, J.; Feuerriegel, S.; Neumann, D. (2017): Setting the Right Tone: How Data Science Enables Investor Communication to Choose the Right Language, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1318-1332

The aforementioned systems partially originate from Information Systems (IS) research. IS research has made significant contributions to develop decision support systems for financial markets, especially for investors [6]. However, this is in contrast to the sending-side of relevant information for financial markets, typically investor relations departments.[1] This group typically lacks customized systems to support the process of preparing financial disclosures. Thus, they cannot rely on such a variety of decision support systems in order to tailor their news flow to the changing needs of capital markets [1].

Academic research has identified two key criteria characterizing high-quality investor communication: a positive tone and a high text readability [7–9]. In fact, many IS studies focus on the relationship between investors and textual information. Most of this research evaluates the tone of language in written communication as an indicator for its impact on investors [6], [9], [10]. Speaking broadly, one can refer to this tone of messages as *sentiment* which is a popular research topic in IS. However, as our understanding of tone of language in a financial market includes more than just the sentiment, namely the impact of language on the stock market, we refer to it as *tonality analysis* in the following. In contrast to sentiment analysis, our approach algorithmically computes the (positive or negative) perception of words to support authors in enhancing the tone. Needless to say, tonality and sentiment analysis are closely related, while the underlying methods can vary considerably [11].

In addition to a positive tonality, academic research reveals that companies must also communicate their messages clearly and concisely in order to reach investor attention. To measure the efficiency of information transmission in messages, readability measures are the quantitative lever of choice [8], [7].

This paper fills the gap of a lacking decision support systems for issuers of financial disclosures. Evidence-based support systems are of high practical value for corporate communication executives to facilitate, e.g. investor communication. This paper therefore suggests an IS prototype which builds on academic knowledge in the area of tonality and readability analysis for capital market communication. Accordingly, our IS prototype works as a decision support system that highlights words with a positive or negative tonality and then provides optional replacements with a better tone for the purpose of improving the readability and tonality of written communication[2]. In order to maximize usability, we pursue a rigorous requirements engineering with potential users prior to programming the IS prototype [12]. Finally, we successfully back-test the enhancements to tonality and readability of our IS prototype in a survey with 66 financial professionals.

The remainder of this paper is organized as follows: Section 2 describes the requirements engineering to identify the user needs. Based on these requirements, Section 3 outlines the textual analysis, while Section 4 describes our user interface

---

[1] Investor relations departments are in charge of communicating the performance of a stock-listed company towards financial markets. For investors, the news flow of a company is a highly relevant information source. Thus the way in which companies compose their investor updates, in line with existing regulations, acts as an important driver of share price movements.

[2] Regulatory disclosures must meet specific requirements regarding the content. However, they are typically not regulated in terms of actual word choice.

including available features. In Section 5, we present the results of validating the tonality enhancement of our IS prototype with financial professionals. Section 6 highlights the managerial implications of our IS prototype. Finally, Section 7 concludes and provides a research outlook.

2 Requirements Engineering

The fact that virtually all stock-listed companies have investor relations and corporate communications departments indicates the importance of communicating a company's performance to external stakeholders, including investors. However, practical knowledge of how investors perceive this communication is scarce as interviews with capital market communication practitioners reveal. Building on [12], we conduct a total of 18 semi-structured interviews with investor relations and communications departments of four global pharmaceutical companies headquartered in the UK, the US and Switzerland and with the investor relations associations of Germany, the UK and Austria.

During the semi-structured interviews, we collect unmet needs and requirements and cluster these into similar themes (e.g. integration into existing processes and simple-to-use interface belong to one such cluster). The interviewed practitioners particularly highlight that they currently lack a decision support tool to guide them how to write financial disclosures in line with the increasing automation of language processing by (automated) investors. Overall, we identify six requirement clusters (two non-functional, four functional), including the following two non-functional requirements:

- *(i) Objective fact base of investors' language perception.* To date, practitioners lack an objective fact base of how investors perceive the language in financial disclosures and process it. Yet, academic research shows that the tone of language matters to investors [7–9].
- *(ii) Defense mechanism against algorithmic trading.* The increasing importance of algorithmic trading, accounting for more than half of the market order volume [13], concerns investor relations practitioners. They particularly claim a misbalance with an increase in algorithmic trading based on language cues in financial disclosures [5], [14], while they do not dispose of tools to adapt their communication to automated trading based on the language used in corporate disclosures. Thus, making knowledge on how investors process language accessible would better equip investor relations against algorithmic trading.

In addition, we identify four functional requirements:

- *(iii) Compliance with insider trading law.* To ensure compliance with insider trading laws, the IS prototype must operate securely and reliably such that accessing confidential data is impossible. Hence, the IS prototype may not include any protocol that tracks information contained in a yet-to-be-published stock-market-relevant announcement. In addition, the tool should maintain its full functionality in the offline mode, e.g. to enable usage during travel.

- *(iv) Simplicity and integration into existing processes.* Practitioners stress the importance of a tool's simplicity and high usability and its seamless integration into existing writing processes. Writing processes shall not become complicated due to yet another tool, but rather be simplified with the help of a user interface that integrates directly into conventional text editors.[3]
- *(v) Collaboration-fostering design.* The process of drafting written corporate communication typically involves several departments (e.g. communications, legal, finance, media). To allow for cross-departmental collaboration, the tool needs to have a self-explanatory reporting dashboard. The dashboard further must be sharable via e-mail or as a PDF.
- *(vi) User-controlled change process.* Corporate disclosures are carefully written based on the available facts, legal requirements and corporate standards. Thus, all changes need to be initiated manually by users, to assure that the IS prototype does not unintentionally modify texts.

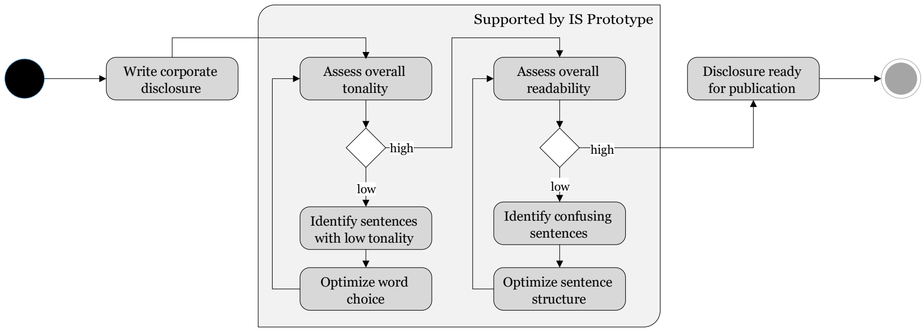


Figure 1. Activity diagram in UML notation showing the reference use case of our IS prototype.

Thus, the IS prototype is intended to support companies in achieving these targets. Moreover, it also incorporates the fact that the editing process of corporate disclosures is highly iterative. The reference use case that serves as the basis to guide our programming efforts and the design of the software system is visualized in a unified modeling language (UML) diagram in Figure 1.[4] As highlighted in the UML diagram, the initial step for companies is to write a first version of the corporate disclosure. Then, the IS prototype helps to assess the overall tonality. In the case of an unsatisfactory tonality, the IS prototype supports users in optimizing their language. In a second step, the IS prototype assesses the overall readability and highlights sentences with

[3] According to our interviews, the predominant text editor among corporate communications departments is Microsoft Word.
[4] The unified modeling language is a common standard to describe the design of a software system. See [15] for details.

optimization potential in case the readability is unsatisfactory. Once the tonality and readability are at satisfactory levels, the disclosure is ready for publication.

Next, we define the end-to-end process behind the development of our IS prototype (see Figure 2). After deriving the requirements of our IS prototype, we collect and pre-process a database of historic financial news disclosures. Then, we extract words that are statistically relevant to investors [11]. This results in a tailored tonality dictionary that contains positive and negative words with relative weights. Finally, we integrate this dictionary into a text editor that fits into existing writing processes. It is noteworthy that we entirely separate our text analytics from the user interface to bundle all computation-intensive analytics in previous processing steps. This ensures that users immediately obtain the correct tonality and readability scores. In the following sections, we further detail the way we generate tonality dictionaries and measure readability.

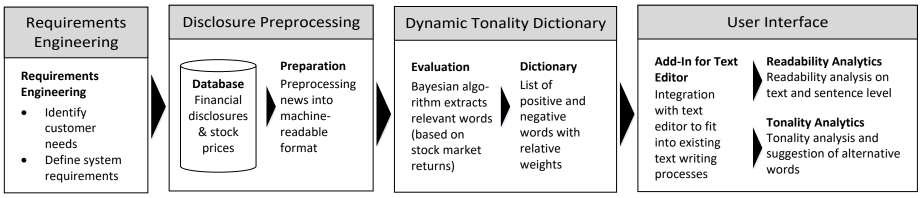


Figure 2. The above flow diagram shows the IS prototype development as an end-to-end process.

3 Data Analytics to Enable Evidence-Based Communication

3.1 Tonality Dictionary Generation

While financial sentiment analysis in IS research primarily counts the relative frequency of positive versus negative words based on standard dictionaries [9], [10], [16], tonality analysis identifies only words that have a relevant and causal effect on buy-or-sell decisions of investors. Among others, this *statistical* approach overcomes a methodological weakness of classical sentiment analysis which typically relies on *manually-selected* word lists that are not adjusted to domain-specific particularities.

For this purpose, we rely on a Bayesian learning approach from [11] that generates tailored dictionaries for the financial domain. Consistent with [17], the authors propose the so-called least absolute shrinkage and selection operator (LASSO) that entails several benefits for dictionary generation. First, this method features a variable selection property that filters out non-informative noise terms and thus, allows to pick only decisive variables in a regression model. Furthermore, the LASSO mitigates the issue of multicollinearity as present when estimating via ordinary least squares. Finally, the LASSO solves the problem of overfitting by finding a reasonable trade-off between bias and variance, which occurs if the model complexity is too high [18]. This results in parsimonious and interpretable models that are particularly suited to extract the words that are most relevant to investors.

As shown in Table 1, we utilize the LASSO approach to generate tailored tonality dictionaries for two different markets along with their corresponding disclosure regulations. First, we cover the U. S. market using a dataset of SEC regulated 8-K filings. Second, we employ a corpus of regulated ad hoc announcements written in English to generate an adjusted dictionary for the German market. In both cases, we measure the stock market reaction following the publication of each financial disclosure. Here, we use the common event study methodology that allows to assess the isolated effect of the financial disclosure on the value of a firm [19], [20].

Table 1. Utilized news sources to generate tailored tonality dictionaries for capital market communication.

Focus	*Market*	*Dataset*	*Benchmark*
Capital market communication	U.S. market	61,241 8-K filings from the years 2004 to 2013	Stock market return
Capital market communication	German market	14,463 ad hoc announcements from the years 2004 to 2011	Stock market return

Based on the input data, i.e. the financial disclosures and the stock market returns, the variable selection property of the LASSO chooses a subset of words that are statistically relevant to investors on the stock market. At the same time, non-informative noise terms are automatically discarded. The coefficient estimates of the LASSO then yields an individual tonality score $s_t \in [-\infty, +\infty]$ for each relevant word.

The resulting weighted word lists, i.e. the tonality dictionaries, allow to statistically assess the overall tonality of financial disclosures. For this purpose, we determine the tonality of a document d by calculating the tonality score $T_d = \sum_t s_t f_{t,d}$. Here, the frequencies $f_{t,d}$ describe how often term t occurs in the document d. The tonality metric T_d thus anticipates the market reaction to the language used in this document. In a next step, the dictionary in combination with the tonality metric can be used to enhance the perceived tonality of a document. For instance, one can exchange words with a negative tonality score with more positive replacements.

Compared to the state of the art, the presented statistical approach significantly outperforms frequently-utilized manually selected positive and negative word lists in terms of explanatory power and predictive performance [11]. Moreover, the generated tonality dictionaries are highly adjusted to domain-specific particularities in capital market communication. In contrast to existing dictionaries that postulate, by definition, an equal importance of all included words, the tonality dictionary also reflects a continuous bandwidth of tonality levels. However, the tonality dictionary only contains words which are used in a financial context and are found to have a statistical impact on the stock market price. Very rare words or words that are not generally used in financial language may be marked as neutral even if these may actually have an impact on reader perception.

3.2 Readability Analytics

An extensive body of literature finds that lower levels of readability reduce investors' reactions to news content [8], [7], [21–24]. Since investors face an ongoing and constant flow of new information, an easy-to-read representation of new information is essential to assure that investors can easily extract relevant information from new financial disclosures. Regulatory handbooks, such as the SEC Plain English Handbook [25], guide the editing of financial disclosures to a certain extent. However, the focus of such regulations is on what to communicate and only to a lesser extent on how to communicate the facts leaving significant freedom to communicators.

In order to fulfill the functional requirements of simplicity and a seamless integration into existing processes, we choose the readability metric based on these criteria. We implement the Automated Readability Index (ARI), as defined in [26], [27], to measure the readability of texts evaluated with our IS prototype. The ARI evaluates the readability of a written document by decomposing sentences into two structural components: the number of words in a sentence and the number of characters per word. The ARI is widely used in readability research [28–30]. In detail, the ARI's formula is defined by

$$ARI = 4.71\,\frac{\text{characters}}{\text{words}} + 0.5\,\frac{\text{words}}{\text{sentences}} - 21.43. \tag{1}$$

Other approaches, i.e., the Gunning fog index, rely on counting the number of syllables per word. However, this requires higher computational effort and results in a worsened error rate in the computation of the index. Thus, we follow [31] and choose ARI as a superior approach.

The readability information is used by practitioners to identify sentences with especially bad readability and to compare the readability of various documents to each other. Sentences with low readability may then be improved by either shortening the sentence or using shorter words within the sentence. While ARI provides a comprehensive and quick assessment of readability, it fails to assess more subtle dimensions of readability such as the usualness of certain words.

4 User Interface

As a result of our requirements engineering, the IS prototype should seamlessly integrate into the existing writing process to promote user adoption. We have therefore developed an IS prototype as an add-in for a conventional text editor, i.e. Microsoft Word. We select this text editor due to its predominance in corporate communication departments, but our IS prototype could be integrated into any other text editor as well (requirement (iv) in Section 2).

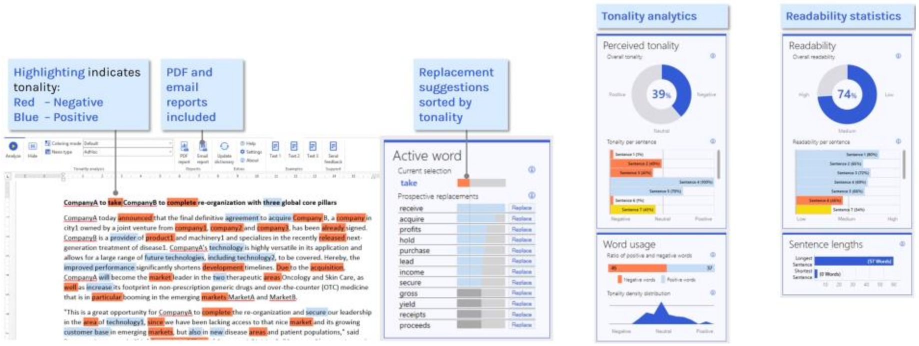


Figure 3. This screen-shot illustrates an exemplary text analysis with our add-in. The word *take*, with a negative tonality, is actively selected by a mouse click. As a result, the active selection panel suggests alternative words, such as the more positive and suitable *acquire*. The center shows the tonality analytics. The right view shows the readability dashboard.

Figure 3 presents our user interface. A drop down menu in the ribbon bar allows users to select a suitable dictionary for the purpose of the underlying communication. For instance, the current implementation of the IS prototype addresses the following domains: (a) a tonality dictionary for capital market communication in the U.S. and the German market following [11], (b) a domain-specific, static dictionary for capital-market communication based on previous research in [16]. This dictionary feature is the core of our application and fulfils the requirements (i) and (ii). The user interface consists of an add-in for a text editor with several functions. As a result of the requirements engineering, the menu (ribbon bar) contains few, but targeted, features. Features include the option to generate a PDF and e-mail report for collaboration (requirement (v)) and a dictionary update function. The dictionary update function allows users to regularly download updated dictionaries from a central server. For this purpose, the add-in automatically communicates with the server and checks for potential dictionary updates. In case updates are available, SSL encryption provides a secure way to transfer the data between server and add-in in line with requirement (iii). To ensure full offline functionality without the need for a continuous internet connection, the dictionaries are encrypted and stored locally after transmission.

Once a suitable dictionary is chosen, users are empowered to objectively analyze the tonality and readability throughout the whole writing process. The tonality analysis highlights words that investors perceive either negatively (highlighted in red) or positively (highlighted in blue). Our IS prototype suggests alternative words as a replacement[5] in case users want to modify the tonality of the underlying document (left

[5] The replacement function builds on the Thesaurus API of Microsoft Word. We perform a refinement to decide which words to display in the add-in. First, we use a part-of-speech tagger to obtain only context-related words. Then we order the prospective replacements according to their tonality score (from most positive to most negative). In the future, we plan to exclude words that are inappropriate in financial communication, i.e. words that not occur at least once in the financial disclosures corpus.

view of Figure 3). In order to ensure that users retain full control of the textual content, word replacements are never applied to the document automatically. Thus, users are required to manually decide whether individual suggestions are adequate and semantically meaningful. This also implies that users are still fully responsible to comply with the regulatory requirements as requested by requirement (vi).

An analytical dashboard displays the aggregated results of the overall tonality (Figure 3, center) and readability (Figure 3, right). The dashboard further reports the tonality and readability of each sentence to simplify the identification of text parts with a particular tonality or readability improvement potential. The dashboard is equipped with full interactivity. As such, a click on a sentence in the vertical bar diagrams lets the editor jump to the corresponding position and highlights the corresponding sentence in yellow color. In addition, we show a density estimation of the tonality scores t_i. A user can easily interpret this plot and verify if the content matches the actual polarity of the content. For instance, one can avoid the case where negative words describe positive information – by that threatening financial markets where readers and automated traders might interpret the information in a wrong way.

To validate and fine-tune our IS prototype, we collect feedback from demo users. The demo users are 49 investor relations and corporate communications professionals from 37 different companies.[6] We provided each demo user with a 2-minutes video tutorial to explain how the IS prototype may be used. We collected feedback on the installation, usability and benefits via a structured online survey. The results of the survey served to fine-tune the IS prototype.

5 Validation of the IS Prototype Through an Expert Study

One recurring concern of corporate communication professionals regarding our data-driven approach is the real-world validity of our algorithmically derived tonality and readability metrics. We encountered skepticism in various conversations with demo users of our IS prototype, despite that fact that we calculate our tonality dictionaries based on real world financial news. Furthermore, the algorithms only include words in the tonality dictionary which have a statistically significant impact on stock prices. As we do not only strive for statistical significance of our tonality dictionaries, but also for economic relevance, we back-test the impact of modifying texts based on our tonality algorithm and readability index with the typical audience of such communications.

We design a survey to gauge the tonality response of financial professionals to stock-market relevant news. The survey consists of eight questions, which cover some of the most frequent topics in financial disclosures [32], such as earnings reports or acquisition activities. In each question, we show an original extract of a published, regulatory filing

[6] The 49 demo users are 18 investor relations officers, 12 investor relations advisors (e.g. from financial communication agencies), 5 heads of investor relations, 4 business translators, 3 heads of media relations and communications, 2 capital market legal advisors, 2 consultants, 1 chief finance officer, 1 investor relations journalist and 1 investment fund manager.

published by stock-listed companies. We keep the original extract in its published form and only disguise the name of the respective company.

Seven questions validate the tonality metric, one question validates the readability metric. Among the seven tonality-related questions, five questions relate to text extracts with negative factual content and two questions relate to text extracts with a positive factual content. In each question, we alter either the tonality or the readability based on our tonality dictionary (cf. Section 3.1) and the ARI readability score (cf. Section 3.2). Thus, for each of the eight questions, we have a pair consisting of the original text extract and a text extract in which we enhance the tonality or readability with the user interface.[7] We ensure that the example remains semantically unchanged. As the thesaurus contains also semantically deviating replacement suggestions, the interface requires the active validation of the replacements by the user.

We conducted the survey among financial professionals between February 3, 2016 and April 4, 2016 in collaboration with a corporate communications service provider. We sent the survey link to more than 4,100 analysts and financial professionals by e-mail. In total, 70 financial professionals participated in the survey of which 66 completed the survey.[8] Across our eight questions, we collect a total of 511 observations, thereof 448 observations for the seven tonality-related questions and 63 observations for the readability-related question. The 66 participants left 17 (out of a total 528) questions unanswered. Overall, the results of the back-testing provide strong evidence for the practical validity of the tonality algorithm and the readability statistics used in the IS prototype.

Finding 1: *Financial professionals evaluate tonality-altered disclosures more positively.* In 70.8 (29.2) percent of the observations, survey participants provide a more favorable 'buy-' or 'sell-'recommendation for the tonality-altered (original) disclosure. This provides strong evidence that the tonality enhancement leads to a more favorable recommendation by financial professionals at a statistically significant level (*t*-statistics 6.064; *P*-Value < 0.001).

Finding 2: *Financial professionals are more sensitive to a better tone in negative than in positive news.* A more detailed analysis reveals that financial professionals react more sensitive to the negative tone in a disclosure with a negative underlying factual news content ('sell' recommendation case) than to the positive tone in a disclosure with positive underlying factual news content ('buy' recommendation case). For the five disclosures with a negative factual content, 76.0 (24.0) percent of the survey participants associate a stronger (weaker) 'Sell' recommendation with the original (tonality-altered) disclosure (*t*-statistics 10.897; *P*-Value < 0.001). For the two disclosures with a positive factual content, 57.5 (42.5) percent of the survey

[7] For negative factual news, we ask *Please select the statement that you associate with a stronger sell recommendation (contains information that is perceived as more negative).* For positive factual news, we ask *Please select the statement that you associate with a stronger buy recommendation (contains information that is perceived as more positive).* For readability, we provide the following question *Please select the statement perceived as easier to read.*

[8] Among the 66 participants, 11 are buy- or sell-side analysts, 7 in investment banking, 5 in communications, 13 in consulting and 30 in other finance-related areas, e.g. advisory. Participants could enter a draw of two Amazon vouchers worth 25 euros each.

participants associate a stronger (weaker) ‘buy’ recommendation with the tonality-altered (original) disclosure (*t*-statistics 1.698; *P*-Value < 0.05). This suggests that the effect of a more positive tone is stronger in negative disclosures. This finding is in line with the psychological negativity bias, which stipulates that human agents weigh negative stimuli more than positive stimuli [33], [34]. Thus, according to the negativity bias, negative information is weighted more heavily than positive information [35].

Finding 3: *Financial professionals rate the readability-altered disclosure as easier to read.* 77.8 percent of the survey participants rate the readability-altered disclosure as easier to read at a statistically significant level (*t*-statistics 5.261; *P*-Value < 0.001).

Table 2. Original and tonality- and readability-altered text pairs of back-testing validation survey (tonality-manipulated words highlighted in bold).

Focus	*Original text extract*	*Altered text extract*
Tonality (neg. factual content)	For the full year 2015, we now **expect** somewhat **weaker** growth for the global economy as well as global industrial and chemical production than was **forecasted** six months ago.	For the full year 2015, we now **anticipate** somewhat **retarded** growth for the global economy as well as global industrial and chemical production than was **predicted** six months ago.
Tonality (neg. factual content)	The oil price-related **decrease** in the Oil & Gas segment **dampened** earnings, while the Functional Materials & Solutions and Chemicals segments provided support through greater contributions.	The Functional Materials & Solutions and Chemicals segments provided support through greater contributions, while the oil price-related **headwinds** in the Oil & Gas segment **compromised** earnings.
Tonality (neg. factual content)	**Contrary** to the general market **expectation**, steel prices **continued to decline**, mainly **due to** a **further** decrease in quotations for Chinese steel exports.	**Different** to the general market **consensus**, steel prices **further fell**, mainly **caused by** an **additional** decrease in quotations for Chinese steel exports.
Tonality (neg. factual content)	The Q3 earnings before interest and taxes (EBIT) of approx. EUR 30 million **turned out significantly weaker than anticipated**.	The Q3 earnings before interest and taxes (EBIT) of approx. EUR 30 million **were behind plan**.
Tonality (neg. factual content)	The company said Thursday that currency fluctuations will **reduce** full-year profit by 85 million pounds, trimming an earlier **forecast** of 100 mio.	The company said Thursday that currency fluctuations will **lower** full-year profit by 85 million pounds, trimming an earlier **prediction** of 100 mio.
Tonality (pos. factual content)	Brew Co. **reported** a better-than-**expected** 7 percent **rise** in underlying third-quarter sales on Thursday, **helped by gains** in Africa and South America, and stronger growth in Europe, where unseasonably mild weather boosted demand.	Brew Co. **published** a better-than-**anticipated** 7 percent **increase** in underlying third-quarter sales on Thursday. **Gains** in Africa and South America, and stronger growth in Europe, where mild weather boosted demand, **stimulated the sales increase**.
Tonality (pos. factual content)	**Following** the transaction, UBM **expects** to **maintain** its investment grade status with the credit rating agencies.	**Succeeding** the transaction, UBM **anticipates** to **confirm** its investment grade status with the credit rating agencies.
Readability	This will be aggravated by losses of sales that are now expected and the consequently missing earnings contributions: the Company’s business prospects are being affected by the worsened economic conditions in China and Russia, which will probably lead to reduced demand, while U.S. business with the commercial vehicle industry as well as the amount of sales to some customers will likely not be at the previously budgeted level	This will be increased by reductions of sales that are now anticipated and the resulting lower earnings contributions. The Company’s business prospects are being affected by the impaired economic conditions in China and Russia, which will probably lead to lessened demand. The U.S. business with the commercial vehicle industry as well as the amount of sales to some customers will likely not be at the previously budgeted level.

Overall, the findings provide supportive evidence for the practical validity of our algorithm-based decision support system for capital market communication. One limitation of our approach is the low number of observations.

6 Managerial Implications

Our IS prototype offers several managerial implications for corporate communication executives. In particular, we have identified three particularly relevant managerial implications:

Implication 1: *Practitioner-friendly tonality analysis.* While current state-of-the-art tonality analysis requires in-depth methodological knowledge as well as advanced analytical skills, the proposed add-in enables any communications practitioner to perform state-of-the-art tonality analysis of disclosures before publication. The add-in thus gives practitioners an information advantage as the perception of texts by investors is forecasted immediately. Furthermore, practitioners are supported with replacement suggestions to improve the tonality and a holistic readability assessment.

Implication 2: *Quality compliance check.* Our IS prototype facilitates a natural compliance check in the writing process as overly negative or positive language is immediately flagged by the tonality assessment. Similarly, issues in the readability become visible when checking the readability metrics of a text. While regulatory frameworks, e.g. the SEC disclosure regulation, provide clear guidance on the content of stock-market-relevant announcements, they only give little guidance on how to write disclosures beyond using plain language [22] [25], [36]. Furthermore, a study of KPMG among investment professionals [1] finds that existing initiatives, such as the SEC Plain English Handbook, have an almost negligible impact in current industry practice. In fact, only 19 percent of the study participants report a positive impact of the plain English initiative on facilitating the disclosure process. In that context, our IS prototype could serve practitioners as a compliance check against regulatory frameworks and support practitioners in putting such regulatory frameworks into practice.

Implication 3: *Process efficiency gains.* The IS prototype integrates directly into conventional text writing processes and offers an objective assessment of the readability and tonality of texts. These assessment results can be easily shared with team members or other departments via the reporting features. Thus, the IS prototype contributes to accelerate the text writing process, especially in settings with multiple stakeholders (e.g. media relations, investor relations, legal, top management). In one interview, we have learned that disclosures of less than 400 words sometimes involve more than 100 iterations. Our tool addresses this process inefficiency by structuring discussions around individual word choices and the overall tone and readability that a company now can deliberately choose in its communication.

In addition, anecdotal evidence from pilot users of the IS prototype confirms the above implications 1 and 3. The director of investor relations at a UK-based pharmaceutical company shares his early user experience in an investor relations professional journal [37]: *"The add-in equips us with deep knowledge on the anticipated reaction of investors. We can now replace subjective feelings by actual evidence when discussing different wording for the very same story"*.

The overall contribution of this IS prototype is twofold: first, we open new avenues for making IS research accessible to practitioners. As such, the prototype serves as a statistically-validated tool to improve form and style of written corporate communication. Second, based on a collaborative approach, the insights and knowledge generated by the practitioners shall flow back to academia, which may then draw conclusions regarding the economics and value of IS on the usage of such a decision support system in practice.

7 Conclusion

In order to successfully compete for investor attention in light of ever-increasing information flows, companies need to ensure that their written communication conveys the relevant information in an easy-to-read manner. Previous research has identified readability and the tone of language as two key elements of successful and positive communication. However, as interviews with investor relations practitioners show, practical tools to steer the investor-specific tone and readability in capital markets communications are not yet available.

The results of the study suggest that improving the tonality of disclosures on basis of the tonality dictionaries leads to significantly more favorable evaluations of the disclosures by financial professionals. In summary, the prototype is a decision support system that not only describes and analyzes data, but directly suggests improvements to users to enhance the tone and readability whilst preparing financial market communication.

References

1. Iannaconi, T.E., Sinnett, W.M.: Disclosure Overload and Complexity: Hidden in Plain Sight. KPMG, Delaware (2011)
2. Laud, R.L., Schepers Donald H.: Beyond Transparency. Information Overload and a Model for Intelligibility. Business and Society Review 114, 365–391 (2009)
3. Financial Reporting Faculty (ed.): Financial Reporting Disclosures: A Programme for Reform. London (2013)
4. Healy, P.M., Palepu, K.G.: Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature. Journal of Accounting and Economics 31, 405–440 (2001)
5. Peterson, R.L.: Markets are a Function of Humans Behavior. The Investor's Brain. Berenberg Equity Highlights 12, 22–26 (2016)

6. Schumaker, R.P., Chen, H.: Textual Analysis of Stock Market Prediction using Breaking Financial News. ACM Trans. Inf. Syst. 27, 1–19 (2009)
7. Rennekamp, K.: Processing Fluency and Investors' Reactions to Disclosure Readability. Journal of Accounting Research 50, 1319–1354 (2012)
8. Tan, H.T., Ying Wang Elaine, Zhuo, B.O.: When the Use of Positive Language Backfires: The Joint Effect of Tone, Readability, and Investor Sophistication on Earnings Judgments. Journal of Accounting Research 52, 273–302 (2014)
9. Tetlock, P.C.: Giving Content to Investor Sentiment. The Role of Media in the Stock Market. The Journal of Finance 62, 1139–1168 (2007)
10. Antweiler, Werner, Frank, Z., M.: Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. The Journal of Finance 59 (2004)
11. Pröllochs, N., Feuerriegel, S., Neumann, D.: Generating Domain-Specific Dictionaries Using Bayesian Learning. In: 23rd European Conference on Information Systems (ECIS 2015). Münster, Germany (2015)
12. Alfano, S., Pröllochs, N., Feuerriegel, S., Neumann, D.: Say It Right: IS Prototype to Enable Evidence-Based Communication Using Big Data. Annals of Information Systems (forthcoming) (2016)
13. Hendershott, T., Riordan, R.: Algorithmic Trading and the Market for Liquidity. J. Financ. Quant. Anal. 48, 1001–1024 (2013)
14. Groß-Klußmann, A., Hautsch, N.: When machines read the news. Using automated text analytics to quantify high frequency news-implied market reactions. Journal of Empirical Finance 18, 321–340 (2011)
15. Dumas, M., ter Hofstede, Arthur H.M.: UML Activity Diagrams as a Workflow Specification Language. In: Gogolla, M., Kobryn, C. (eds.) UML 2001. The Unified Modeling Language : Modeling Languages, Concepts, and Tools : 4th International Conference, Toronto, Canada, October 1-5, 2001 : Proceedings, pp. 76–90. Springer, New York (2001)
16. Loughran, T., McDonald, B.: IPO First-Day Returns, Offer Price Revisions, Volatility, and Form S-1 Language. Journal of Financial Economics 109, 307–326 (2013)
17. Taddy, M.: Distributed Multinomial Regression. Ann. Appl. Stat. 9, 1394–1414 (2015)
18. Tibshirani, R.: Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society 58, 267–288 (1996)
19. Konchitchki, Y., O'Leary, D.E.: Event Study Methodologies in Information Systems Research. International Journal of Accounting Information Systems 12, 99–115 (2011)
20. MacKinlay, C.A.: Event Studies in Economics and Finance. Journal of Economic Literature 35, 13–39 (1997)
21. Henry, E.: Are Investors Influenced By How Earnings Press Releases Are Written? Journal of Business Communication 45, 363–407 (2008)
22. Loughran, T., McDonald, B.: Regulation and Financial Disclosure. The Impact of Plain English. J Regul Econ (Journal of Regulatory Economics) 45, 94–113 (2014)

23. Rogers, J.L., van Buskirk, A., Zechman, S.L.C.: Disclosure Tone and Shareholder Litigation. The Accounting Review 86, 2155–2183 (2011)
24. Tan, H.T., Wang, E.Y., Zhou, B.: How Does Readability Influence Investors' Judgments? Consistency of Benchmark Performance Matters. The Accounting Review 90, 371–393 (2015)
25. Securities and Exchange Commission (SEC): A Plain English Handbook: How to Create Clear SEC Disclosure, http://www.sec.gov/pdf/handbook.pdf
26. Senter R. J., Smith E. A.: Automated Readability Index (1967)
27. Kincaid, J.P., Fishburne, Jr , Robert P, Rogers, R.L., Chissom, B.S.: Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel (1975)
28. Ghose, A., Ipeirotis, P.G.: Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics. IEEE Trans. Knowl. Data Eng. 23, 1498–1512 (2011)
29. Klare, G.R.: Readable computer documentation. ACM J. Comput. Doc. 24, 148–168 (2000)
30. Moffit, K., Burns, M.B.: What Does That Mean? Investigating Obfuscation and Readability Cues as Indicators of Deception in Fraudulent Financial Reports. In: Proceedings of the 15th American Conference of Information Systems (AMCIS 2009). Association for Information Systems (2009)
31. Hu, N., Bose, I., Koh, N.S., Liu, L.: Manipulation of Online Reviews: An Analysis of Ratings, Readability, and Sentiments. Decision Support Systems 52, 674–684 (2012)
32. Feuerriegel, S., Ratku, A., Neumann, D.: Which News Disclosures Matter? News Reception Compared Across Topics Extracted from the Latent Dirichlet Allocation. SSRN Journal (2015)
33. Kanouse, D., Hanson, R.: Negativity in Evaluations. In: Attribution. Perceiving the Causes of Behaviour, pp. 47–62. General Learning Press, Morristown (1972)
34. Kanouse, D.E.: Explaining Negativity Biases in Evaluation and Choice Behavior. Theory and Research. Advances in Consumer Research 11, 703–708 (1984)
35. Peeters, G., Czapinski, J.: Positive-Negative Asymmetry in Evaluations: The Distinction Between Affective and Informational Negativity Effects. European Review of Social Psychology 1, 33–60 (1990)
36. Shi, C., Pukthuanthong, K., Walker, T.: Does Disclosure Regulation Work? Evidence from International IPO Markets. Contemporary Accounting Research 30, 356–387 (2013)
37. Alfano, S., Feuerriegel, S., Neumann, D.: Compliance in IR. Informed, 40–41 (2016)

SENSEMAKING AND COMMUNICATION ROLES IN SOCIAL MEDIA CRISIS COMMUNICATION

Stefan Stieglitz[1], Milad Mirbabaie[1], Lara Schwenner[1], Julian Marx[1], Janina Lehr[1], Felix Brünker[1]

University of Duisburg-Essen, Department of Computer Science and Applied Cognitive Science, Duisburg, Germany
{stefan.stieglitz, milad.mirbabaie, lara.schwenner, julian.marx, janina.lehr, felix.brünker} @uni-due.de

Abstract. Social media is becoming increasingly important during crisis situations. Affected people are now enabled to provide helpful on-site information, and emergency service organisations can use social media to inform people and communicate with them. This study addresses how different communication roles in social media affect sensemaking during crises situations. To this end, we conducted a study on Twitter during the Brussels attacks of 2016. We collected a sample of 3,223,197 tweets, which included a total of 1,535,943 participants. Our study reveals that, whereas information distribution dominates early crisis stages, attention-keeping gains in importance in subsequent stages. It is decisively depending on the characteristics and retweet behaviour of certain communication roles that information is being consulted by individuals in a situation of either lack of information or information overload.

Keywords: Sensemaking, Social Media, Twitter, Crisis Communication, Information Systems

1 Introduction

Crisis situations are characterised by ambiguity, confusion and feelings of disorientation. Thus, during disasters, people have a powerful occasion for sensemaking [1] to rationalise what is going on [2]. While sensemaking is the process of social construction by information seeking, sensegiving is a process by which attempts are made to influence the meaning construction and sensemaking of others towards a preferred interpretation of an occurrence [3]. People have a drive for sensemaking due to the motivation of simplification and the desire to construe the world in favourable ways [4]. From one of its first definitions, sensemaking is a process of social construction that occurs especially at that point, when discrepant cues interrupt individuals' ongoing activity in their real life. Thus, sensemaking involves the retrospective development of plausible meanings that rationalise what people are doing and making sense of a situation after it took place [2, 5]. To start

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Stieglitz, S.; Mirbabaie, M.; Schwenner, L.; Marx, J.; Lehr, J.; Brünker, F. (2017): Sensemaking and Communication Roles in Social Media Crisis Communication, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1333-1347

their sensemaking process, individuals need information to connect different cues and draw a picture of the crisis situation.

Social media services such as Facebook or Twitter are increasingly used to communicate during crisis situations by individuals, members of the affected public, professional media and organisations [6, 7]. The microblogging service Twitter especially emerged as a widely used social reporting tool to instantly spread information on social crises [8]. Recent research shows that Twitter is a resilient and rapid information diffusion tool under large-scale crises such as natural disasters [9–11], terror attacks [12, 13] or social movements [14, 15]. Due to its short texting service interface on cell phones, Twitter turned out to be extremely rapid in tweeting situational reports to the online community [16], thus allowing first responders to collectively cope with the crisis situation [17]. During large-scale crises, it has become the norm for the incident to be initially reported by a local eyewitness with a mobile communication device. The report is rapidly distributed through social media services, and mainstream media involvement follows [12, 18]. However, especially in social media communication during large-scale crises, there might either occur the problem of information dearth or the problem of information overload, depending on the particular crisis situation [19]. Therefore, besides a general need for information, people also try to select incoming information in social media. Both the provision of information as well as information selection can be undertaken by hashtags [20], but also through opinion leaders in a social network [21]. Twitter users not only consume the incoming information from their network, but also broadcast the consumed information into their own network [22]. By that, the most active users might direct the public sensemaking process in a social media crisis communication by information selection and providing. In this context, recent research mainly focused on communication patterns during crises [22].

Recent research primarily focused on face-to-face situations [4, 23], though less on sensemaking processes in social media environments. Social media is used as an information source in crisis situations, because it provides fast and easy access [24]. On the one hand, there is a need for information when it comes to natural disasters or terror attacks to start the sensemaking process; on the other hand, communication systems, such as social media applications, might be overloaded with information [19]. Thus, information tends to be chaotic and uncontrollable and it is difficult to maintain a vast overview. This can cause rumours or false information in dangerous situations and might hinder emergency services to manage the crisis efficiently.

An Information-Systems-Journal article [20] identified collective sensemaking through Twitter during the 2011 Egypt Revolution. It could be monitored as the emergence of few hashtags out of many, which mainly brought together crisis-related information. Conclusively, the authors revealed that – besides hashtags – one can investigate whether power users, who receive the highest number of retweets by others [25], direct the collective sensemaking process. If the existence of power users can be confirmed, one can further analyse them and their characteristics. In this paper, we build on the conclusions and suggestions of [20] and address the collective sensemaking and its dynamics through roles and characteristics in a case study. We therefore aim to answer two research questions:

1. Which roles can be identified within the collective sensemaking process in social media during a crisis situation and how do they develop over time?
2. Which characteristics do these roles adopt in the collective sensemaking process in social media?

The remainder of this paper proceeds as follows: in Section 2, we present the status quo of the literature regarding (1) crisis communication in social media, and (2) sensemaking in crisis situations. In Section 3, we introduce our research design, which includes the case description, the data collection and our data analysis in detail. We then present the results of our case study in Section 4, followed by a discussion of our research findings in Section 5. The paper ends with a conclusion in Section 6, including limitations and an outlook for further research.

2 Theoretical Background

Crises are characterised by high levels of threat, situational uncertainty and decision-making pressure under time constraints. Therefore, a critical need for immediate and accurate information occurs in crisis situations, which are usually provided by experts, emergency management professionals, governmental organisations or similar authority figures. Besides the face-to-face-context and traditional media, in the past few years, social media use has also become a consistent feature in crisis response [7, 19, 26]. In a crisis communication process, social media tools are being more and more used by each actor of society, like individuals or media, but also increasingly by formal crisis responders [6, 7, 27]. People use and rely on official sources and other believable eyewitness accounts from which to source their information [28]. In this context [29], six types of information resources under different crises types can be distinguished on Twitter: eyewitness, government, non-governmental organisation, business, traditional and/or internet media and outsiders. While tweets from governments mainly advice or warn the public, tweets from the media offer information about crisis development, whereas outsiders produce information, although they are not personally affected by the event. Depending on the role type, the social media usage in a crisis communication can differ: individuals especially tend to use social media during disasters and post-disasters to investigate what is going on, check with family and friends or mostly direct, relay, synthesize or redistribute (existing) information [9]. Nevertheless, organisations (e.g., emergency management agencies like the police) use social network to spread important information to the public by using microblogging channels for two-way interaction [30–32]. This behaviour was observed during the Queensland Floods in 2010/2011 [33, 34], in the information-sharing behaviour of the US police departments [35] and during the Boston Marathon bombings of 2013 [36]. EMA use their social media channels mainly to broadcast accurate and simple messages to keep populations informed [31], [37]. Furthermore, [37] mention that a government's communication behaviour on social media is dependent on who is tweeting: the PR-department (formal, one-way-communication) or an employee (informal two-way communication). Especially during social crises, companies need to spread reliable information as early as possible to avoid rumours in social media crisis communication [17]. Overall, social media technologies mediate human communications in social crisis situations and present varying patterns of crisis communication [14]. In the social media context, e.g., in Twitter, users not only consume the incoming information from their own network, but also diffuse information to their own network. Furthermore, [21] showed through conducting an analysis of a 260 million Twitter dataset, that approximately 20,000 elite users (celebrities, media, organisations and bloggers) were responsible for 50 percent of the tweets. The authors suggest that people get their news directly through the elite users, who produce information, as well as through intermediaries (ordinary, non-elite users), which have a high follower count and distribute incoming information to their network [21]. Also, user characteristics, for example, the user's popularity, might be a driver for information diffusion [38], as well as URL and hashtag inclusion [39]. A social media crisis communication might involve users who seek information to start their sensemaking process, but also the ones who perform as sensegivers by information providing and sharing. Besides elite users and opinion leaders, in a case study about implementing an online platform for crisis communication, [40] suggest three different user types, which might play an important role for sensemaking in a social media communication: (1) inspectors, who define a certain event, (2) contributors, who provide media content and witness statements, and finally, (3) investigators, who verify media content by organising and sorting data and detecting missing information. In contrast, [41] distinguish the helper, reporter, retweeter and repeater as active user types on Twitter – each with their specific characteristic of producing, distributing and organising information.

Especially in crisis situations, people become highly suggestible and turn to others to find cues [42]. Collective sensemaking is manifested as the communication behaviours of active information seeking, offering and sharing among a like-minded group of people, which helps reduce the level of situational ambiguity and collectively defines an unfolding situation [20, 43]. Collective sensemaking can take place among emergent groups of actors, who interpret information together face-to-face [44], or remotely through social media [35, 45]. Through their ability to facilitate collective sensemaking, social media serves the purpose of filling in the possible information vacuum left by mainstream media [19] or other official channels. Crisis communication differs across media types like Facebook and Twitter [46–50], and therefore, the sensemaking process can also differ. But independent of the type of media and the leading roles, the direction of the sensemaking process is also dependent on the structure of a social network. During the 2011 Egypt Revolution, [20] detected the occurrence of collective sensemaking through collecting information and maintaining situational awareness via hashtags. In detail, the authors revealed that the Twitter space was structured around a few hashtags out of many, which can be related to the keynoting phenomenon [42]. Furthermore, during the 2011 Egypt Revolution, the hashtags' frequency changed over time and a content analysis showed that hashtags were used either as a symbol (to focus attention to a certain issue) or as a word in a sentence (to distribute information to a certain issue). Thus, [20] suggest that there are two characteristics of collective sensemaking through Twitter: (1)

maintaining a high level of awareness to a certain issue during an unstable situation (attention keeping), and (2) offering situational news and information about temporal events (information distributing). The authors conclude that, besides hashtags, one can investigate whether power users direct the collective sensemaking process and how these power users are characterised. In this context, power users could be defined as Twitter users, who receive the highest number of retweets [22], because retweet frequency can be seen as a measure of popularity for the message or its author [51].

3 Research design

In order to address our research questions, we conducted a case study and focused on the Brussels attacks in March 2016 as a type of a public crisis which generated a significant amount of attention and traffic on social media. We have chosen the microblogging platform Twitter as our data source, because recent research shows that Twitter as a social media application is frequently used for crisis communication in social crises or social change [14, 17, 20, 52]. Especially for the 2016 Brussels attacks, there is a direct link between Twitter and the crisis itself: in the first few hours after the bombings, the Crisis Center Belgium (@CrisiscenterBE) called on the people to communicate via social network sites, because mobile communication networks collapsed temporarily.

3.1 Case description

On the morning of 22nd of March 2016, three coordinated nail bombings occurred in Belgium: two at the Brussels Airport in Zaventem, and one at the Maalbeek metro station in Brussels. In the scope of the attacks, a total of 32 victims were killed. Furthermore, over 300 people were injured. Belgium raised the terror threat level to its highest, the public and air traffic of Brussels was suspended, and the population was told to stay where they were. As it was unclear who was responsible for the attacks and whether the suspects were still alive, the search and investigation went on. Thus, the population remained in uncertainty for nearly ten hours. Shortly after the attacks, the crisis communication began to spread over to social network applications.

3.2 Data collection and analysis

For our empirical analysis, we collected tweets regarding the Brussels attacks 2016 and examined the first 27 hours of the crisis, from March 22nd (8:00 am CET) to March 23rd (10:59:59 am CET), 2016. We collected the data through the Search API[1] of Twitter with a self-developed Java crawler, using the library Twitter4J[2]. We saved the collected data in a MySQL database, gathering all tweets that contained at least one of the following three keywords (including the hashtags): *brussels*,

[1] https://dev.twitter.com/rest/public/search, last access: 08-22-2016
[2] http://twitter4j.org, last access: 08-22-2016

brusselsattacks and *bruxelles* and merged them into one dataset. The keywords were selected based on their frequent usage on the platform. After collecting the data, we decided to divide the data into seven time slots in order to handle the large amount of Twitter data. The slots are of different lengths (slot 1-5 cover three hours, slot 6-7 cover six hours) because of the tweet volume and the beginning dynamics of the communication, which decreases over the examined overall time period.

For identifying the roles of participants in the crisis communication, and thus, to answer our first research question, we used methods from social network analysis. For each time slot, we analysed the graph given by the retweet network. The nodes (vertices) of our network are Twitter accounts and the edges are retweets. The network is therefore a directed network and the edge weights are the number of retweets. To analyse and visualise the networks, we used the open source tool Gephi[3]. To produce the visualisations for each time slot, we ran the layout algorithm ForceAtlas 2. The size and colour of each node represent the number of retweets from a node. Subsequently, we filtered each time slot with the aid of the Gephi filter Giant Component to remove all nodes which are not connected to the main network. Furthermore, we highlight the nodes with the highest indegree with their account names. Indegree is the value of how much a node has been retweeted, whereas the outdegree describes the value of how much a node has retweeted itself.

As an elementary assumption, we suggest that the phenomenon of keynoting is not only applicable to the utilisation of hashtags, but also to the determination of roles during crisis situations – especially under consideration of power users of a social network. The retweet function is one of the core functions of information diffusion on Twitter. The retweet activity not only shows reciprocal relations between different users, but is also suitable for identifying particularly active or strong connected users [53]. The decisive key figure to determine power users by the means of a network analysis using Gephi is the indegree measure. Since the available data is represented in a directed graph, the indegree indicates a node's prestige by quantifying its frequency of being retweeted. The retweet frequency for certain tweet messages can be seen as a measure of popularity for the message or its author [51]. For the purpose of arranging a significant sample of power users, we identified the top 20 power users for each time slot based on their indegree value. We focused on the top 20 users, because these users have the most impact on sensemaking in our case.

We also calculated the overall follower count, overall original tweet count, and betweenness centrality value based on the directed graph. Betweenness centrality measures the degree to which a node is in a position of brokerage by summing up the fractions of shortest paths between other pairs of nodes that pass through it [54]. The resulting set of power users are subsequently categorised into roles by all involved authors independently. Similar to [29], we adopted information sources such as traditional/internet media or the government as roles. In contrast to [29], we extended the source of outsiders to the role of private person and various kind of public persons to gain more detailed insights. The categorisation leads us to seven different role types. To complete the analysis regarding the first research question, identified roles

[3] https://gephi.org, last access: 08-22-2016

will be examined towards their development over time and how the presence of certain roles might shift during several crisis stages.

For the second research question, role characteristics are further examined. At that point, we combined findings from existing literature, in particular the definition of attention keepers and information distributors [20] with metrics of the Twitter network and the results of a detailed content analysis concerning all original tweets posted by the top 3 users of each role type sorted by indegree, hence the number of retweets received by those users. To obtain further insights on the characteristics, we conducted a qualitative content analysis [55], considering all original tweets (131) posted by the top 3 accounts of each role type (ranked by indegree). Indeed, 131 tweets represent less than 10 percent of the total number of tweets in the subdataset, but for two of the seven identified role types, there could be related only three user accounts each in the subdataset. Therefore, we decided to examine a smaller set of user accounts' original tweets, but with the same number of accounts per role (namely 3), instead of examining a larger set, but with a dissimilar number of related user accounts per role type. The 131 original tweets of the present selection by 21 user accounts (7 (roles) x 3 (top users)) were evaluated by all authors independently for the purpose of characterising the function each role adopts in the collective sensemaking process in a crisis situation. This includes the measure hashtag-use (symbolic/word) like [20] did, and furthermore, measures URL-use, which is a common measure to characterise tweet content [39], and the type of information (solicitousness, crisis information, opinion, other), which were partially adopted by [29] and further determined by the authors after exploring a random set of tweets out of the 2016 Brussels bombing dataset. We adopted relevant information types from [29], such as sympathy and emotional support, which we redefined as solicitousness. Furthermore, we applied caution and advice, as well as affected individuals, which we combined with infrastructure and utilities to general crisis information. Since we analysed the first 27 hours, and [29] argue that those are the most relevant information types in the early stages of a crisis, we excluded information types like donations and volunteering. Since neither information type addressed opinion sharing, we introduced another information type. Whether a hashtag is used as a symbol or as a word indicates the affiliation to either: 1) attention keeping or 2) information distribution. URL-use supports the designated characteristic of information distribution [39], whereas characterising the information type facilitates the understanding of role characteristics.

4 Findings

The results are split into two parts. The first part shows the results according to the first research question, including the results of the social network analysis and the analysis of the roles' development over time. The second part presents all relevant results to answer the second research question regarding the role characteristics. In total, we received a sample of 3,223,197 tweets, which includes a total number of 1,535,943 participants. In Table 1, the results for each time slot are summarised.

Table 1. Metrics of the directed network per each slot.

Day	Slot	Period (CET)	Edges	Nodes	Diameter	Giant Component (% of nodes)
22nd of March 2016	1	08:00-10:59:59	32620	33252	5	70.96%
	2	11:00-13:59:59	377116	265631	25	95.52%
	3	14:00-16:59:59	557533	386740	25	92.51%
	4	17:00-19:59:59	521538	383329	23	93.16%
	5	20:00-22:59:59	411192	318198	15	93.07%
23rd of March 2016	6	23:00-04:59:59	316421	234150	14	89.93%
	7	05:00-10:59:59	166365	134923	11	86.22%

First, we conducted a social network analysis to identify the top 20 power users ranked by indegree for each time slot. Figures 1-7 represent the social network analysis, which illustrate the top 20 power users for each time slot.

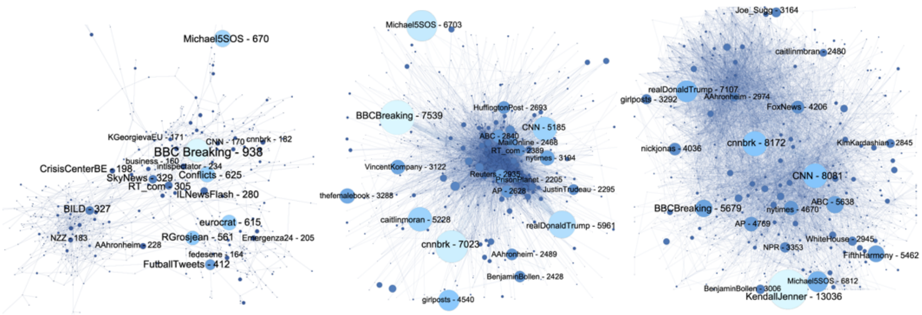


Figure 1. Social Network Analysis results: Time slots 1-3 (from left to right)

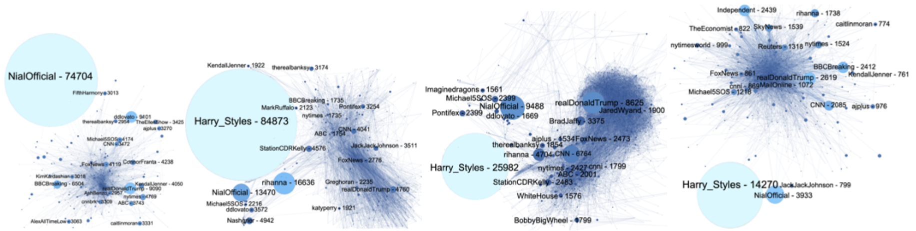


Figure 2. Social Network Analysis results: Time slots 4-7 (from left to right)

By analysing the accounts of the identified top 20 power users (ranked by indegree) of our dataset (table 2), we categorised them in seven distinct roles, derived from [29]: media organisations (e.g.: @CNN); 2) governmental organisations (e.g. @WhiteHouse); 3) private persons; 4) public persons (journalists); 5) public persons (celebrities) (e.g.: @Harry_Styles); 6) public persons (politicians) (e.g.: @RealDonaldTrump); 7) public persons (other) (e.g. @Pontifex).

Table 2. Top 20 users (ranked by overall indegree) categorised into roles

Rank	Role Type	Indegree	Betweenness Centrality (Rank)		Outdegree (Rank)		Follower Count
1	Public Person (celebrity)	125,125	0	(8)	1	(9)	28,6 m
2	Public Person (celebrity)	101,595	0	(8)	1	(9)	25,7 m
3	Public Person (politician)	38,162	0	(8)	4	(7)	9,48 m
4	Media Organisation	29,798	60433.41	(3)	36	(5)	25,57 m
5	Media Organisation	24,806	12677.3	(5)	23	(6)	23,6 m
6	Public Person (celebrity)	24,192	0	(8)	1	(9)	6,23 m
7	Public Person (celebrity)	23,079	0	(8)	1	(9)	62,3 m
8	Public Person (celebrity)	19,008	0	(8)	1	(9)	16,9 m
9	Media Organisation	18,686	41094.0	(4)	23	(6)	39,57 m
10	Media Organisation	18,347	166685.24	(2)	45	(3)	28,01 m
11	Media Organisation	15976	200576928	(1)	59	(1)	6.2 m
12	Media Organisation	14435	6173	(7)	50	(2)	9.4 m
13	Public Person (celebrity)	14342	0	(8)	1	(9)	36.9 m
14	Public Person (journalist)	11813	0	(8)	1	(9)	568 k
15	Public Person (celebrity)	8475	0	(8)	1	(9)	2.9 m
16	Private Person	7982	0	(8)	3	(8)	1.45 m
17	Public Person (other)	7832	0	(8)	1	(9)	6.19 m
18	Media Organisation	7396	6405.5	(6)	38	(4)	7.9 m
19	Public Person (celebrity)	6793	0	(8)	1	(9)	3.28 m
20	Public Person (celebrity)	5863	0	(8)	1	(9)	45.8 m

Second, we manually analysed the development of each role over time (Figure 3). For this step, we identified roles' total number of indegree for each time slot. The overall indegree of each role contains the number of indegree of every power user account, which could be assigned to the specific role.

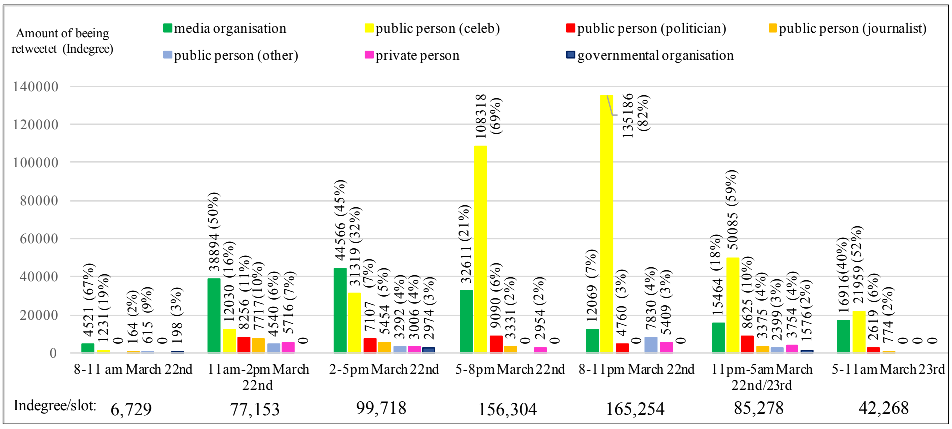


Figure 3. Roles' Development over time by indegree

In the dataset, we could identify two leading roles over time: 'media organisations' dominate over the first three time slots. Then, the role 'public persons' (celebrities) takes the lead until the end of the dataset. In the dataset, we could identify two peaks.

The first peak occurred on 22nd of March through the role media organisation, 2-5pm, the second on 22nd of March, 8-11pm, through the role public persons (celebrities).

As a second consecutive step of analysing upcoming roles in crisis situations, we defined role characteristics for participating roles during the Brussels attacks. Regarding hashtag-use, we found hashtags and keywords used as words within sentences far more often among all role types except governmental organisations (e.g. “All public transport in #Brussels shut down…”). The usage of hashtags as symbols was rather unusual concerning our sample (e.g., “I entrust to God’s mercy all those who lost their lives. #Brussels”). The usage of URLs, which is designated to serve as an indicator for information content, turned out to be non-existent among public persons, whereas media organisations dominated the dissemination of URLs on Twitter. To accomplish the content analysis, the information type of each original tweet was matched to one or more of following categories: 1) crisis information, 2) solicitousness/condolences, 3) opinion, and 4) other (everything what does not fit into category 1-3. Media organisations were found to primarily post crisis information, as well as public persons (journalists) and governmental organisations. Solicitousness and condolences were expressed by public persons (celebrities), private persons and public persons (others). Two-thirds of Tweets stemming from public persons (politicians) contained opinions, the other third contained solicitousness, and the remaining contained content.

Due to a high value in indegree, a high amount of shared crisis information and the usage of hashtags as words, following roles, could be found to be information distributors: media organisations, public persons (journalists) and governmental organisations. To the characteristics of attention keeping, public persons (celebrities), public persons (politicians), private persons and public persons (other) could be matched. As a subordinate characteristic, we defined central users who show a high indegree value and a high betweenness centrality at the same time. From all involved roles, only media organisations suited these special characteristics (see Table 2).

5 Discussion

This study provides findings regarding important sensemaking actors and their different contributions to the sensemaking process, which supports the findings of [20], according to which power users play an integral role during collective sensemaking one way or another. Although retweeting behaviour has various motivations, retweeted messages seem to have a higher informational value than non-retweeted ones [39]. Adopting this theory to power users, their tweets could be perceived with a higher informational value and therefore be selected by individuals for the sensemaking process among a social community. Our findings are consistent with [21], suggesting that people get their news through elite users, who are generally the producer of information. In this context, we identified seven roles, which play a role in the sensemaking process about the Brussels bombings 2016 communication in Twitter. The identified roles are partially consistent with the information sources by [29] regarding media organisations and governmental organisations. In this study, we extended the role of the outsider [29] – people that are not personally involved or affected by the event, to private persons and different kind of public persons like journalists, celebrities, politicians and others. While recent research mainly focused on sensemaking activities regarding information production and distribution [12, 22], our results support the findings of [20], according to which there are two characteristics of the collective sensemaking process. Even though celebrities seem to be important actors during the sensemaking process, their contribution to sensemaking differs from the attributes of media organisations. Media organisations, public persons (journalists) and governmental organisations could be primarily identified as information distributors. Public persons (celebrities or politicians or other), private persons, however, acted as attention keepers in the Brussel crisis communication, maintaining situational awareness by expressing their solicitousness or opinion. To relate these characteristics to the identified role types, we described them in a structural manner based on the Twitter network and the tweeted content. Besides the indegree value, media organisations can be characterised by a high betweenness centrality, a high outdegree, a high URL-use and a word-hashtag-use. In comparison, celebrities are characterised by no betweenness-centrality, a low outdegree, a high follower count, no URL-use and a word-hashtag-use. Contrary to [20], attention keepers in this study predominantly used word-hashtags rather than symbol-hashtags.

The follower count did not seem to predict the user's impact on sensemaking, since, for example, the user with the highest follower count was only on the seventh position in our top 10 ranking by indegree. Still, it is a determining factor for information diffusion, independent of the role type [38] and must be considered by examining online sensemaking. Additionally, a high overall betweenness centrality could be measured for the communication patterns of media organisations, while the betweenness centrality regarding tweets by public persons (celebrities) was nearly non-existent. Taking into account that the social network analysis was conducted for a retweet network, the betweenness centrality can be interpreted as a bridging function within the network that allows information to diffuse on the shortest routes. A non-existent score however emerges from not retweeting other user's content, while a high score identifies users, which are not only central to the social network, but also consumed and broadcasted other user's information. This is of high importance for the flow of crisis communication within a network and therefore for the sensemaking process [20, 43]. Those assumptions are reflected by the top 20 roles' outdegree. Whereas celebrities' indegree arose from one tweet alone, media organisations have an original tweet count varying between 23 and 45 tweets during the first 27 hours. In fact, media organisations posted on average at least one tweet per hour. The findings emphasise the influence of media organisations on public communication in social networks (as gatekeepers) and their role in the collective sensemaking process.

The observation of the role's development during the crisis moreover holds valuable information considering the sensemaking process. Similar to [29], our findings reveal that media organisations seem to be a dominant role during the early stage of a crisis, which can be defined as the event-breakout-phase [22], and is characterised by breaking news rather than local information avenues. Media organisations' high number of retweets can be explained by the public's need for information, which constitutes the starting point for the individual sensemaking process [1]. People turn to social media, since it serves the purpose of filling in the possible information vacuum left by mainstream media and therefore facilitates collective sensemaking [19].

Our findings are consistent with [29] stating that media organisations show the highest fraction of tweets in crises situations that are instantaneous like the Brussels bombings. Simultaneously, large-scale crises on social media can elicit the problem of information dearth or information overload, depending on the particular crisis situation [19]. People need to find a way to select the incoming information in a social media crisis communication, which can be comprehended by focusing on power users (as roles). In both cases, power users can either provide relevant information to fill the information gap or help start people's sensemaking process by sharing relevant content and distance themselves from social media noise. Adopting the suggestion by [42], according to which the sensemaking process consists of the chaotic milling and organised keynoting interactions, our study provides evidence that not only hashtags can facilitate keynoting interactions [20], but also power users e.g. celebrities drawing attention to the crisis by declaring solicitousness, or media organisations distributing information about the development of crisis, or establishing or distributing symbols (e.g. pictures) related to the crisis. These actors might help people overcome the chaotic situation when a crisis breaks out by providing and selecting relevant information. Consistent with [42], during the breakout-phase, media organisations seem to hold the most dominant voices, which emerge to override the chaotic opinions of the crowd (milling) by starting organised keynoting interactions.

However, the dominant sensemaking role of media organisations can only be observed for the early stages of the course of events. After a few hours, their number of retweets decreases, while the number of retweets of celebrities heavily increases. Their role in collective sensemaking in crisis situations has not been respected in recent literature yet, though our study provides evidence that elite users like celebrities contribute strongly to the people's sensemaking process. Celebrities' influence seems to be even greater than information provided by media organisations, since their indegree are significantly higher. Regarding the development of the roles, we observed that in the early stages of a crisis event, people turn to social media to gather and share relevant information for their individual sensemaking process. Roles, which act as information distributors, are dominant sensegivers at this point, as they try to support and direct the sensemaking of others [3]. After the event-breakout-phase, celebrities take over as dominant roles, acting as attention keepers by sharing solicitousness and opinions to influence people's meaning construction [3].

6 Conclusion

In this study, we analysed the Twitter communication of the Brussels attacks (2016). Through conducting a social network analysis, we could identify seven roles which are relevant for the collective sensemaking process during a crisis. Although some roles seemed to have a larger impact on sensemaking than others, every single categorised role contributed to the collective sensemaking process.

Given the complexity of human-involved collective sensemaking, focusing on retweets might not provide a holistic view. As the API is the only access point to large-scale Twitter data available to researchers outside of Twitter as a platform, there is no opportunity to independently verify the quality of the dataset. To mitigate this problem, [56] suggest creating more specific parameter sets with different users, bounding boxes, and keywords. We furthermore argue that a sample of 131 original tweets for content analysis is not large enough to generalise the different role characteristics we found. In the meantime, we analysed a bigger sample, but reached similar results. Nevertheless, for further research, we plan to expand the sample. We also note that our conclusions are case-specific and cannot be generalised without care and without examining a number of further crises, especially of different types.

Following the leading point of sensemaking through roles in crisis situations, the analysis of the case revealed significant aspects of social interaction in crisis communication. We contributed to the understanding of roles and their influence in social media during the sensemaking processes. Furthermore, this is one of the few studies that considers the dynamics of an event. In the underlying analysis, we detected a shift in the dominance of role characteristics regarding consecutive crisis stages. On a temporal layer, role characteristics alter from information distribution in early crisis stages to attention keeping in subsequent stages. The information being consulted by individuals decisively depends on the behaviour of major roles. In fact, information diffusion through retweets is a crucial means for collective sensemaking. If a situation generates a lack of information, (central) roles provide fast access to information, which can be immediately spread within a network or social cluster. In case of an information overload, the activity of roles in social media affects the selection of information to urge the sensemaking process. Through following the flow of information based on roles, the shaping of public opinion can be assessed more precisely.

Based on our findings, we recommend for further research to search for these role types in other Twitter networks automatically. Of course, these structural role descriptions have to be verified with other, larger datasets. Another approach that hasn't been addressed so far is the danger of upcoming rumours during crisis situation. [17] point out the importance of spreading trustworthy information as early as possible to avoid these. Since rumours are also part of the collective sensemaking process, one could examine the roles' impact on upcoming rumours in a social media crisis communication and the perceived trustworthiness by the public.

References

1. Maitlis, S., Sonenshein, S.: Sensemaking in crisis and change: Inspiration and insights from weick (1988). J. Manag. Stud. 47, 551–580 (2010).
2. Weick, K.E.: Sensemaking in Organisations. (1995).
3. Glola, D. a., Chittipeddi, K., Gioia, D.A., Chittipeddi, K.: Sensemaking and Sensegiving in Strategic Change Initiation. Strateg. Manag. J. 12, 433–448 (1991).
4. Chater, N., Loewenstein, G.: The under-appreciated drive for sense-making. J. Econ. Behav. Organ. 126, 137–154 (2016).
5. Weick, K.E., Sutcliffe, K.M., Obstfeld, D.: Organizing and the Process of Sensemaking. Organ. Sci.

16, 409–421 (2005).
6. Hughes, A.L., Palen, L.: The Evolving Role of the Public Information Officer: An Examination of Social Media in Emergency Management. J. Homel. Secur. Emerg. Manag. 9, (2012).
7. Palen, L., Anderson, K.M., Mark, G., Martin, J., Sicker, D., Palmer, M., Grunwald, D.: A vision for technology-mediated support for public participation & assistance in mass emergencies & disasters. Proc. 2010 ACMBCS Visions Comput. Sci. Conf. 1–12 (2010).
8. Jin, Y., Liu, B.F., Austin, L.L.: Examining the Role of Social Media in Effective Crisis Management: The Effects of Crisis Origin, Information Form, and Source on Publics' Crisis Responses. Communic. Res. 41, 74–94 (2011).
9. Vieweg, S., Hughes, A.L., Starbird, K., Palen, L.: Microblogging during two natural hazards events: what twitter may contribute to situational awareness. CHI 2010 Cris. Informatics April 10–15, 2010,. 1079–1088 (2010).
10. Kwon, H.Y., Kang, Y.O.: Risk analysis and visualization for detecting signs of flood disaster in Twitter. Spat. Inf. Res. 24, 127–139 (2016).
11. Jung, C., Tsou, M., Issa, E.: Developing a Real-time Situation Awareness Viewer for Monitoring Disaster Impacts Using Location-Based Social Media Messages in Twitter. Int. Conf. Locat. Soc. Media Data. 1–5 (2015).
12. Oh, O., Agrawal, M., Rao, H.R.: Information control and terrorism: Tracking the Mumbai terrorist attack through twitter. Inf. Syst. Front. 13, 33–43 (2011).
13. Potts, L., Mapes, K.: HCI in Business, Government, and Organizations: Information Systems. 9752, 72–81 (2016).
14. Starbird, K., Palen, L.: (How) will the revolution be retweeted?: information diffusion and the 2011 Egyptian uprising. Proc. acm 2012 Conf. CSCW. 7–16 (2012).
15. LeFebvre, R.K., Armstrong, C.: Grievance-based social movement mobilization in the #Ferguson Twitter storm. New Media Soc. 1–21 (2016).
16. Lerman, K., Ghosh, R., Surachawala, T.: Social Contagion : An Empirical Study of Information Spread on Digg and Twitter Follower Graphs. Proc. Fourth Int. AAAI Conf. Weblogs Soc. Media. 90–97 (2010).
17. Oh, O., Agrawal, M., Rao, H.R.: Community Intelligence and Social Media Services: A Rumor Theoretic Analysis of Tweets During Social Crises. MIS Q. 37, 407–426 (2013).
18. Oh, O., Kwon, K.H., Rao, H.R.: An Exploration of Social Media in Extreme Events: Rumor Theory and Twitter during the Haiti Earthquake 2010. In: ICIS. p. 231 (2010).
19. Shklovski, I., Palen, L., Sutton, J.: Finding community through information and communication technology in disaster response. Proc. 2008 ACM Conf. Comput. Support. Coop. Work. ACM. 127–136 (2008).
20. Oh, O., Eom, C., Rao, H.R.: Research Note —Role of Social Media in Social Change: An Analysis of Collective Sense Making During the 2011 Egypt Revolution. Inf. Syst. Res. 210–223 (2015).
21. Wu, S., Hofman, J.M., Mason, W. a., Watts, D.J.: Who says what to whom on twitter. Proc. 20th Int. Conf. World Wide Web. 705–714 (2011).
22. Kwon, H.K., Oh, O., Agrawal, M., Rao, R.H.: Audience Gatekeeping in the Twitter Service: An Investigation of Tweets about the 2009 Gaza Conflict. AIS Trans. Human-Computer Interact. 4, 212–229 (2012).
23. Bordia, P., Difonzo, N.: Problem Solving in Social Interactions on the Internet: Rumor As Social Cognition. Soc. Psychol. Q. 67, 33–49 (2004).
24. Austin, L., Liu, B.F., Jin, Y.: How Audiences Seek Out Crisis Information: Exploring the Social-Mediated Crisis Communication Model How Audiences Seek Out Crisis Information: Exploring the Social- Mediated Crisis Communication Model. J. Appl. Commun. Res. 40, 188–207 (2012).
25. Oh, O., Tashmasbi, N., Rao, R.H., Vreede, G.-J.: A Sociotechnical Vie of Information Diffusion and Social Changes: From Reprint to Retweet. ICIS 2012 Proc. 1–11 (2012).
26. Heverin, T., Zach, L.: Use of microblogging for collective sense-making during violent crises: A study of three campus shootings. J. Am. Soc. Inf. Sci. Technol. 63, 34–47 (2012).
27. Mirbabaie, M., Ehnis, C., Stieglitz, S., Bunker, D.: Communication roles in public events – A case study on Twitter communication. In: Information Systems and Global Assemblages. (Re)Configuring Actors, Artefacts, Organizations. pp. 207–218 (2014).
28. Starbird, K., Palen, L., Hughes, A.L., Vieweg, S.: Chatter on the red: what hazards threat reveals about the social life of microblogged information. CSCW '10 Proc. 2010 ACM Conf. Comput. Support. Coop. Work. 241–250 (2010).
29. Olteanu, A., Vieweg, S., Castillo, C.: What to Expect When the Unexpected Happens: Social Media Communications Across Crises. Proc. 18th ACM Conf. Comput. Support. Coop. Work Soc. Comput. -

CSCW ’15. 994–1009 (2015).
30. Ehnis, C., Mirbabaie, M., Bunker, D., Stieglitz, S.: The role of social media network participants in extreme events. In: 25th Australian Conference of Information Systems (2014).
31. Sutton, J.N., Johnson, B., Greczek, M., Spiro, E.S., Fitzhugh, S.M., Butts, C.T.: Connected Communications : Network Structures of Official Communications in a Technological Disaster. Proc. 9th Int. ISCRAM Conf. 1–10 (2012).
32. Krüger, N., Stieglitz, S., Potthoff, T.: Brand Communication In Twitter - A Case Study On Adidas. In: PACIS 2012 Proceedings (2012).
33. Bruns, A., Burgess, J.E., Crawford, K., Shaw, F.: #qldfloods and@QPSMedia: Crisis Communication on Twitter in the 2011 South East Queensland Floods, http://eprints.qut.edu.au/48241/, (2012).
34. Ehnis, C., Bunker, D.: Social Media in Disaster Response: Queensland Police Service - Public Engagement During the 2011 Floods. Proc. 23rd Australas. Conf. Inf. Syst. 1–10 (2012).
35. Heverin, T., Zach, L.: Twitter for city police department information sharing. Proc. Am. Soc. Inf. Sci. Technol. 47, 1–7 (2010).
36. Ehnis, C., Bunker, D.: The impact of disaster typology on social media use by emergency service agencies: The case of the Boston marathon bombing. In: 24th Australasian Conference on Information Systems (2013).
37. Bergstrand, F., Landgren, J., Green, V.: Authorities don ’ t tweet , employees do ! In: MobileHCI (2013).
38. Macskassy, S. a, Michelson, M.: Why Do People Retweet ? Anti-Homophily Wins the Day ! Proc. Fifth Int. Conf. Weblogs Soc. Media - ICWSM ’11. 209–216 (2011).
39. Suh, B., Hong, L., Pirolli, P., Chi, E.H.: Want to be retweeted? Large scale analytics on factors impacting retweet in twitter network. In: Proceedings - SocialCom 2010: 2nd IEEE International Conference on Social Computing, PASSAT 2010: 2nd IEEE International Conference on Privacy, Security, Risk and Trust. pp. 177–184 (2010).
40. Blum, J., Kefalidou, G., Houghton, R., Flintham, M., Arunachalam, U., Goulden, M.: Majority report: Citizen empowerment through collaborative sensemaking. ISCRAM 2014 Conf. Proc. - 11th Int. Conf. Inf. Syst. Cris. Response Manag. 767–771 (2014).
41. Reuter, C., Heger, O., Pipek, V.: Combining Real and Virtual Volunteers through Social Media. Iscram 2013. 780–790 (2013).
42. Turner, R.H., Killian, L.M.: Collective Behavior. Prentice Hall College Div; 3 Sub edition (1987).
43. Dailey, D., Starbird, K.: “It’s Raining Dispersants.” Proc. 18th ACM - CSCW’15 Companion. 155–158 (2015).
44. Kendra, J., Wachtendorf, T.: The Waterbourne Evacuation of Lower Manhatten on September 11: A case of Distributed Sensemaking. (2006).
45. Vieweg, S., Palen, L., Liu, S.B., Hughes, A.L., Sutton, J.: Collective Intelligence in Disaster : Examination of the Phenomenon in the Aftermath of the 2007 Virginia Tech Shooting. Iscram. 44–54 (2008).
46. Hughes, A.L., St. Denis, L. a. a., Palen, L., Anderson, K.M.: Online public communications by police & fire services during the 2012 Hurricane Sandy. Proc. 32nd Annu. ACM Conf. Hum. factors Comput. Syst. - CHI ’14. 1505–1514 (2014).
47. Kaufhold, M.A., Reuter, C.: The Self-Organization of Digital Volunteers across Social Media: The Case of the 2013 European Floods in Germany. J. Homel. Secur. Emerg. Manag. 13, 137–166 (2016).
48. Birkbak, A.: Crystallizations in the Blizzard : Contrasting Informal Emergency Collaboration In Facebook Groups. Proc. Nord. Conf. Human-Computer Interact. 428–437 (2012).
49. Bruns, A., Stieglitz, S.: Quantitative Approaches to Comparing Communication Patterns on Twitter. J. Technol. Hum. Serv. 30, 160–185 (2012).
50. Bruns, A., Stieglitz, S.: Twitter Data: What Do They Represent? it - Inf. Technol. 56, 240–245 (2014).
51. Kwak, H., Lee, C., Park, H., Moon, S.: What is Twitter , a Social Network or a News Media? Int. World Wide Web Conf. Comm. 1–10 (2010).
52. Stieglitz, S., Bruns, A., Krüger, N.: Enterprise-Related Crisis Communication on Twitter. Proc. der 12. Int. Tagung Wirtschaftsinformatik (WI 2015). 917–932 (2015).
53. Ahn, H., Park, J.-H.: The structural effects of the sharing function on Twitter networks: Focusing on the retweet function. J. Inf. Sci. 41, 354–365 (2015).
54. Brandes, U.: On variants of shortest-path betweenness centrality and their generic computation. Soc. Networks. 30, 136–145 (2008).
55. Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. (2008).
56. Morstatter, F., Pfeffer, J., Liu, H., Carley, K.: Is the sample good enough? Comparing data from Twitter’s streaming API with Twitter’s firehose. Proc. ICWSM. 400–408 (2013).

Ermittlung des Virtualisierungspotenzials von Beratungsleistungen im Consulting

Volker Nissen[1], Henry Seifert[2]

[1] TU Ilmenau, FG WI für Dienstleistungen, 98684 Ilmenau, Deutschland
volker.nissen@tu-ilmenau.de
[2] Mieschke Hofmann und Partner, Gesellschaft für Management- und IT-Beratung mbH, Königsallee 49, 71638 Ludwigsburg, Deutschland
henry.seifert@mhp.com

Abstract. Die Beratungsbranche sieht sich, genau wie ihre Klienten, ständig neuen Herausforderungen und sich verändernden Rahmenbedingungen gegenüber. Beratungsanbieter sollten daher ihr Leistungsportfolio immer wieder kritisch infrage stellen. Obwohl sie die Wettbewerbsfähigkeit ihrer Klienten durch innovative Lösungen stärken und dort maßgeblich an der Entwicklung neuer Konzepte zur Digitalisierung beteiligt sind, wird bei der Erbringung von Beratungsleistungen oft nur auf traditionelle Face-to-Face Ansätze zurückgegriffen. Ein virtueller Prozess ist demgegenüber ein Prozess, in dem die physische Interaktion verschwindet. Der Übergang eines physischen Prozesses hin zu einem virtuellen Prozess wird als „Prozess Virtualisierung" bezeichnet. Virtualisierung ist ein Trend, dem sich Beratungsunternehmen auch hinsichtlich ihrer eigenen Geschäftsprozesse stellen müssen. Theoriegeleitet sowie auf Basis ergänzender empirischer Forschung leiten wir in diesem Beitrag ein mögliches Vorgehen ab, um das Virtualisierungspotenzial von Beratungsprozessen (oder deren Teilschritten) im konkreten Fall ex ante beurteilen zu können.

Keywords: Virtualisierung, Digitale Transformation, Unternehmensberatung, Geschäftsmodellinnovation, Beratungsprozess, Beratungsforschung.

1 Grundlagen und Motivation

Unternehmensberatung (Consulting) kann definiert werden als professionelle Dienstleistung, die durch eine oder mehrere, im allgemeinen fachlich dazu befähigte und von den beratenen Klienten hierarchisch unabhängige Person(en) zeitlich befristet sowie meist gegen Entgelt erbracht wird [18]. Sie hat zum Ziel, betriebswirtschaftliche Probleme des beauftragenden Unternehmens interaktiv mit den Klienten zu definieren, strukturieren und analysieren, sowie Problemlösungen zu erarbeiten, und auf Wunsch ihre Umsetzung gemeinsam mit Vertretern des Klienten zu planen und im Unternehmen zu realisieren [18]. In diesem üblichen Verständnis ist Unternehmensberatung ein „people business". Dabei suchen Berater Klienten auf, um

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Nissen, V.; Seifert, H. (2017): Ermittlung des Virtualisierungspotenzials von Beratungsleistungen im Consulting, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1348-1362

mit ihnen gemeinsam vor Ort interaktiv (Face-to-Face) unternehmensrelevante Problemstellungen zu bearbeiten.

Die Beratungsbranche ist, trotz positiver Umsatzentwicklung, mit neuen Herausforderungen konfrontiert. Einerseits nimmt die Konkurrenz mit Anbietern aus Billiglohnländern und Freelancern im Bereich von Standardleistungen zu. Daneben betreten neue Wettbewerber mit innovativen Geschäftsmodellen und technologiegetriebenen Beratungsansätzen den Markt. Auf der Klientenseite ist eine zunehmende Professionalisierung im Einkauf und Umgang mit Unternehmensberatung zu beobachten. Auch hat die Preis-Sensitivität zugenommen.

Die Virtualisierung von Beratungsleistungen kann, angesichts der beschriebenen Herausforderungen, eine innovative Strategie zur Sicherung des nachhaltigen Unternehmenserfolgs sein und klassische Angebote der Unternehmensberatung ergänzen. Virtualisierung zielt darauf ab, den Anteil direkter Face-to-Face Interaktion zwischen Berater und Klient durch den geeigneten Einsatz von Informations- und Kommunikationstechnologien (IKT) zu reduzieren [11] [20]. Sie kann damit als Strategie zur digitalen Transformation des Beratungsgeschäftes bezeichnet werden.

Notwendig zur Realisierung dieser ambitionierten Vision sind geeignete Artefakte, wie webbasierte Beratungsplattformen, mobile Beratungsapplikationen, semantische Technologien, Data Mining- und Kooperations-Werkzeuge. Ebenso werden Konzepte benötigt, wie die Virtualisierbarkeit von Beratungsprozessen beurteilt und Lösungen, die konventionelle und virtualisierte Beratung zu einem Ganzen verbinden, entwickelt werden können. Methoden und Techniken der Virtualisierung sind zu schaffen. Hier besteht noch erheblicher Forschungsbedarf [26].

Während die raumzeitliche Flexibilität bei der Virtualisierung zunimmt, kann die reduzierte persönliche Interaktion von Beratern und Klienten deren Beziehung beeinträchtigen. Ob Virtualisierung der richtige Ansatz ist, einen Beratungsprozess zu verändern und wie dies am besten geschehen kann, muss intensiv geprüft werden. Hier besteht heute jedoch ein Mangel an belastbaren Kriterien und Virtualisierungskonzepten. Daraus leitet sich die Forschungsfrage des vorliegenden Beitrags ab:
Wie kann das Potenzial zur Virtualisierung einer Beratungsleistung (oder eines Teilschrittes) anhand konkreter Kriterien ex ante beurteilt werden?

2 Inhaltliche Einordnung der Arbeit (Related Work)

Die hier bearbeitete Forschungsfrage ist dem Consulting Research [18] [19] zuzurechnen, worunter die wissenschaftliche Beschäftigung mit der Dienstleistung Unternehmensberatung, den Beratungsunternehmen als Organisationen und dem Beratungsmarkt mit seinen verschiedenen Teilnehmern auf Anbieter- und Nachfragerseite verstanden wird. Consulting Research hat zwei Anliegen [18]. Erstens, die wissenschaftliche Durchdringung des Themas Unternehmensberatung, wobei der von einzelnen Beratungsprojekten abstrahierende wissenschaftliche Erkenntnisgewinn im Mittelpunkt steht. Zweitens, die Übertragung wissenschaftlicher Theorien, Erkenntnisse und Methoden auf die unternehmerische Praxis mit dem Ziel, Aufgabenstellungen und Probleme im Umfeld von Beratungsprozessen und

Beratungsunternehmen besser als bisher zu lösen. Der vorliegende Beitrag ordnet sich schwerpunktmäßig beim zweiten Kernanliegen des Consulting Research ein, da letztlich eine Vorgehensweise entwickelt wird, die Beratungsanbieter bei der digitalen Transformation ihres Geschäftes unterstützen soll.

Gleichzeitig hat unser Beitrag eine gewisse Nähe zum Thema Service Modularisierung [7] [12] [25]. Durch die gezielte Anwendung von Standardisierung und Modularisierung wird die Virtualisierung von Einzelaufgaben und Teilprozessen in der Beratung vorbereitet. Diese Module können, gegebenenfalls bedarfsgerecht neu kombiniert, die Basis für innovative Beratungsleistungen bilden.

3 Methodik und Datengrundlage

3.1 Überblick

Zur Beantwortung der Forschungsfrage wird im ersten Schritt auf die theoretische Grundlage der *process virtualization theory* von Overby [23] [24] zurückgegriffen und diese in die Domäne Consulting übertragen. Ergänzend liefert eine großzahlige Studie zu den Einflussfaktoren des Virtualisierungspotenzials im deutschen Beratungsmarkt weitere Hinweise für ein sinnvolles Vorgehen. Im dritten Schritt wird eine strategische Perspektive eingenommen, die zusätzlich Chancen und Risiken der Virtualisierung auf Basis einer literaturgestützen Delphi-Studie mit Beratern und Klienten einbezieht. Im Ergebnis ergibt sich ein Analyseprozess in drei Schritten, der im weiteren Verlauf des Beitrages genauer dargestellt wird. Die methodischen Grundlagen dazu werden nachfolgend kurz dargelegt.

3.2 Prozessvirtualisierung – theoretische Fundierung

Ein virtueller Prozess ist ein Prozess, in dem die physische Interaktion zwischen den Menschen und/oder Maschinen verschwindet. Der Übergang eines physischen Prozesses hin zu einem virtuellen Prozess wird als „Prozess Virtualisierung“ bezeichnet. Overby [23] [24] entwickelte die generisch angelegte ‚Process Virtualization Theory‘ (PVT) und führte den Begriff der „Prozess Virtualisierbarkeit“ ein. Overby sieht die Nutzung/Akzeptanz und die Qualität der Prozessergebnisse als Basis, um die Virtualisierbarkeit eines Prozesses (als abhängige Größe) ex post messen zu können. Daneben benennt er eine Reihe von Einflussgrößen, die sich auf die Virtualisierbarkeit eines Prozesses auswirken und demnach für die hier angestrebte ex ante Beurteilung des Virtualisierungspotenzials von Beratungsprozessen grundsätzlich geeignet erscheinen. Die PVT erscheint als der passende theoretische Anker für diese Arbeit, weil sie im Gegensatz zur Theorie des Task-Technology-Fit nicht der Evaluation der Eignung einer Technologie für eine bestimmte Aufgabe dient, sondern vielmehr der Erklärung, wie gut sich eine Aufgabe (als Teil eines Prozesses) prinzipiell für den Technologieeinsatz eignet. Sie ist hier auch dem Technology-Acceptance-Model (TAM) vorzuziehen, da Beratungsleistungen als Co-Kreation von Berater und Klient primär aus Prozessperspektive (und nicht im Sinne reiner Technologieakzeptanz) zu betrachten sind [24]. Auch wurde das TAM bezüglich seiner mangelnden Relevanz und Nützlichkeit, zukünftige Systeme besser zu gestalten, stark kritisiert [2].

Gemäß der PVT wirken sich insbesondere die folgenden Prozesseigenschaften negativ auf die Virtualisierbarkeit eines Prozesses aus: hohe sensorische Anforderungen (da physische Interaktion entfällt), hohe Anforderungen an die persönliche Beziehungsebene der Beteiligten und daran anknüpfende Konstrukte wie Vertrauen, hohe Anforderungen an die Synchronität von Aktivitäten in der Prozessausführung und hohe Identifikations- und Steuerungsanforderungen im Prozess (da bei Virtualisierung die tatsächlich Interagierenden leichter verschleiert werden können).

Demgegenüber mildern die folgenden Eigenschaften des (IT-basierten) Virtualisierungsmechanismus die genannten Prozesscharakteristika und wirken sich somit auf die Virtualisierbarkeit eines Prozesses aus: die Fähigkeit der IT prozessrelevante Informationen darzustellen („representation"), die Fähigkeit der IT eine zeit- und ortsunabhängige Prozessteilnahme zu ermöglichen („reach"), sowie die Möglichkeiten durch IT eine Authentifikation der Prozessbeteiligten und Überwachung des Prozessablaufes zu gewährleisten („monitoring capability").

Balci und Rosenkranz [3] merken an, dass die Messbarkeit der Prozess-Virtualisierbarkeit bisher kaum empirisch untersucht ist. Sie finden in einer eigenen Untersuchung empirische Bestätigung für die PVT, sehen jedoch gleichzeitig Hinweise auf deren Unvollständigkeit. Demnach spielen auch Faktoren eine Rolle, die Merkmale der Prozessteilnehmer (z.B. IT-Kenntnisse) betreffen.

3.3 Kriterien der Virtualisierbarkeit – empirische Studie

Der Aspekt einer möglichen Unvollständigkeit der PVT wird für den Anwendungsbereich Consulting durch eine ergänzende großzahlige Befragung zu den Kriterien der Virtualisierbarkeit von Beratungsleistungen aufgegriffen. Befragt wurden, in Kooperation mit dem Bundesverband Deutscher Unternehmensberater BDU e.V., Gesellschaften aller Größen aus der gesamten Consultingbranche.

Im Rahmen der hier beschriebenen Studie wurden Daten online mit Hilfe der Umfragesoftware *Unipark QuestBack* erhoben. Mit Hilfe einer Online-Befragung ergeben sich zahlreiche Vorteile, wie die geringen Erhebungskosten, der Entfall der manuellen Dateneingabe oder die schnelle Verfügbarkeit der Daten [1]. Den Vorteilen bei Online-Befragungen stehen allerdings auch Nachteile gegenüber [13]. So entsteht aufgrund der Anonymität im Internet beispielsweise ein erhöhtes Risiko von Mehrfachteilnahmen. Dieser Aspekt wurde u.a. dadurch minimiert, dass der BDU aus seinen Kontakten Einladungen zur Teilnahme an die Mitglieder versendete. Da der BDU bezüglich seiner Mitglieder einen guten Querschnitt der deutschen Beratungslandschaft darstellt, sind hier Verzerrungen des Ergebnisses kaum wahrscheinlich. Weiterhin wurde viel Wert auf die Verständlichkeit des Fragebogens gelegt und dies im Rahmen eines Pre-Tests verifiziert.

Die Befragung der Teilnehmer wurde im Zeitraum vom 23. November bis 18. Dezember 2015 durchgeführt. Insgesamt klickten 765 Teilnehmer auf den Link zur Startseite der Online-Befragung, der zuvor in einer E-Mail versendet wurde. 654

Teilnehmer starteten die eigentliche Online-Befragung. In der anschließenden Editierung und Kodierung der Daten fand eine Datenbereinigung statt. Dabei wurden 102 Antworten ausgeschlossen, da sie die Anforderungen an Vollständigkeit nicht erfüllten. In Summe konnten so 552 Fragebögen für die weitere statistische Analyse berücksichtigt werden. Insgesamt wird der deutsche Beratungsmarkt nach Umsatz und Beratungsfeldern gut abgebildet, so dass von weitgehend repräsentativen Ergebnissen ausgegangen werden kann. Auch hinsichtlich Alter und Berufserfahrung der Befragten zeigte sich, dass das Meinungsbild von Beratern mit unterschiedlichen Erfahrungsniveaus erfasst werden konnte. Die Repräsentativität der Stichprobe ist insofern einzuschränken, weil die Datenerhebungsmethode bei allen Vorzügen nicht ganz unproblematisch. So setzt diese Methode voraus, dass Teilnehmer in der Lage sind, den Online-Fragebogen entsprechend zu nutzen. Teilnehmer, die wenig Internet-affin sind, könnten hier ausgegrenzt werden [13], was bei der Zielgruppe der Unternehmensberater jedoch kaum wahrscheinlich ist.

Die Befragung zielte insgesamt darauf ab, den Status Quo und die Perspektiven der digitalen Transformation in der Unternehmensberatung in Deutschland zu klären. In diesem Beitrag sollen jedoch nur Kriterien betrachtet werden, um das Virtualisierungspotenzial von Beratungsleistungen zu bestimmen. Hierzu wurde vorab eine strukturierte Literaturanalyse nach Webster und Watson [27] durchgeführt, um mögliche Kandidaten für solche Kriterien zu identifizieren. Neben der originären Literatur zur Beratungsvirtualisierung wurde auch in den Themenbereichen Telearbeit, Telekooperation, Computer Supported * Work, Task Analysis, E-Government und E-Services recherchiert, um auf das Consulting potenziell übertragbare Resultate zu identifizieren. Weiterhin sind Beiträge, die von Overby zitiert wurden oder die Overby zitieren, berücksichtigt worden. Insgesamt konnten 41 Beiträge für die Recherche genutzt werden. Durch die anschließende Literatur-Synthese und Erstellung einer Konzeptmatrix wurde eine Liste an Kriterien der Virtualisierbarkeit erarbeitet. Diese Kriterien lassen sich in die drei Gruppen *Klient*, *Beratung* und *Beratungsaufgabe* clustern. Die besondere Relevanz dieser drei Gruppen bei der Untersuchung der Virtualisierung von Beratungsleistungen wird in der Literatur von Autoren wie Wurdack [28] herausgestellt. Darüber hinaus wurden diese Gruppierungen bereits im Kontext der Standardisierung von Beratungsleistungen genutzt [10] und auch im Bereich der Modularisierung von Dienstleistungen als Schlüsselaspekte benannt [7].

Die Gruppe *Klient* beinhaltet alle Kriterien, die im Zusammenhang mit der Integration des Klienten stehen. Dies umfasst sowohl das Vertrauen des Klienten in das Beratungsunternehmen als auch die Akzeptanz des Klienten für die virtuelle Beratungsleistung. Ferner können die technischen Anforderungen des Klienten sowie die Erfahrung des Klienten mit virtuellen Beratungsleistungen genannt werden.

Die zweite Gruppe *Beratung* beinhaltet alle Kriterien, die im Zusammenhang mit der Beratungsorganisation, also dem Beratungsunternehmen selbst, stehen. Hierzu zählen sowohl die Erfahrung der Beratung mit virtuellen Beratungsleistungen als auch die Reife des Wissensmanagements. Weiterhin sind die Auslastung der Berater und die Seniorität der Berater bestimmend für die konkrete Virtualisierbarkeit.

Die dritte Gruppe *Beratungsaufgabe* deckt die Kriterien ab, die eine Evaluation des Virtualisierungspotenzials auf Ebene der Beratungsaufgabe ermöglichen. Dies umfasst sowohl die Kritikalität, d. h. das Risiko und Konfliktpotenzial der Aufgabe, als auch die Komplexität und mithin die Veränderlichkeit, Vielzahl und Vielfalt der Aufgabe. Ferner gilt es, die Dringlichkeit, also den Termindruck der Aufgabe sowie die Wichtigkeit, dass bedeutet die individuelle Bedeutung der Aufgabe, zu bestimmen. Weitere Kriterien, die in diese dritte Gruppe gehören, sind die Vertraulichkeit der Informationen und Aufgabe und die Interaktivität, d. h. die Frequenz, Dauer, Art und Intensität d. Berater-Klienten-Interaktion. Ein wesentliches und abschließendes Kriterium der Gruppe Beratungsaufgabe ist die Individualität und mithin der Anspruch des Klienten an eine individuelle Lösung.

Nach der literaturgestützten Herleitung dieses initialen Kriterienkatalogs galt es, die Kriterien im Rahmen einer großzahligen Befragung von Unternehmensberatern zu evaluieren. Um das Virtualisierungspotenzial praxistauglich beurteilen zu können, ging es hierbei auch darum, die Anzahl der bislang 15 Kriterien auf ein in den Unternehmen handhabbares Maß zu reduzieren.

Die Ergebnisse der Literaturanalyse bildeten die Grundlage einer entsprechenden Frage in der Online-Studie, bei der die Teilnehmer einerseits die Relevanz dieser Kriterien anhand einer 6-stufigen Likert-Skala beurteilen sollten, andererseits aber auch die Freitext-Möglichkeit hatten, weitere Kriterien zu ergänzen und zu beurteilen. Insgesamt konnten in unserer Studie hierzu die Aussagen von 374 Teilnehmern ausgewertet werden.

Um die Kriterien der Virtualisierbarkeit von Beratungsleistungen zu charakterisieren, bietet sich die Faktorenanalyse (EFA) als dimensionsreduzierendes Verfahren an. Hierdurch können latente Faktoren identifiziert werden, die entsprechend fokussiert die Einflussfaktoren wiedergeben [9]. In diesem Zusammenhang wurde für die 15 initialen Kriterienkandidaten auch Cronbach's Alpha berechnet, das grundsätzlich Werte zwischen 0 und 1 annehmen kann. Je höher der Wert von Cronbach's Alpha, desto höher ist die Korrelation zwischen den Indikatoren und damit die Interne-Konsistenz-Reliabilität [8]. Dementsprechend sollte ein Cronbach's Alpha Wert $\geq 0{,}7$ vorliegen [22]. Bei uns beträgt dieser Wert 0,83.

Für die Durchführung der EFA wurde zunächst der Datensatz mittels einer Complete-Case Analyse bereinigt und auf seine prinzipielle Eignung für die Analyse untersucht. Zentrale Kriterien, die darüber Aufschluss geben, ob ein Datensatz für eine EFA grundsätzlich geeignet ist, sind das KMO Kriterium und der Bartlett-Test [4]. Hierbei deutet ein KMO-Wert von über 0,5 daraufhin [15], dass die Daten grundsätzlich eine gewisse Korrelation aufweisen und damit für eine EFA genutzt werden können. Im vorliegenden Fall beträgt der KMO-Wert 0,81. Der Bartlett-Test deutet seinerseits bei einer Ablehnung der Nullhypothese auf eine grundsätzliche Eignung der Daten für eine EFA hin. Für die vorhandenen Daten ist dieser signifikant ($p < 0{,}001$) von Null verschieden. Die Nullhypothese, dass die Korrelationsmatrix nur zufällig von der Einheitsmatrix verschieden ist, kann also abgelehnt werden, womit auch der Bartlett-Test die Eignung der Daten für eine EFA bestätigt.

Der Tabelle 1 ist zu entnehmen, dass sich die Items recht trennscharf den einzelnen Faktoren zuordnen lassen, da jedes Item, mit zwei Ausnahmen, bei einem cutoff-Wert von 0,5 lediglich auf einen Faktor lädt. Gleichzeitig liegen bei den meisten Items Faktorladungen über 0,7 vor. Im Ergebnis lassen sich die 15 Kriterienkandidaten auf 7 Faktoren reduzieren, denen wir aussagekräftige Namen zu ihrem jeweiligen Fokus gegeben haben. Diese bilden nun die wichtigste Grundlage, um die Virtualisierbarkeit von Leistungen der Unternehmensberatung ex ante zu beurteilen.

Tabelle 1. Faktoranalyse zu den Kriterien der Virtualisierbarkeit von Beratungsleistungen.

Rotierte Komponentenmatrix	Komplexität	Verfügbarkeit geeigneter Berater	Akzeptanz d. Klienten	Reife d. Beratungs-organisation	Dringlichkeit d. Beratungsleistung	Vertrauen zwischen d. Akteuren	Interaktivität
Kritikalität	,888						
Komplexität	,856						
Vertraulichkeit	,770						
Individualität	,693						
Seniorität der Berater		,859					
Auslastung der Berater		,833					
Erfahrung des Klienten mit virtuellen Beratungsleistungen			,843				
Anforderungen des Klienten an die zu nutzende Virtualisierungstechnologie			,701				
Akzeptanz des Klienten für die virtuelle Beratungsleistung			,574			(,566)	
Erfahrung der Beratung mit virtuellen Beratungsleistungen				,859			
Reife des Wissensmanagements				,790			
Dringlichkeit					,888		
Wichtigkeit	(,546)				,627		
Vertrauen des Klienten in das Beratungsunternehmen						,896	
Interaktivität							,757

Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Varimax mit Kaiser-Normalisierung. Die Rotation ist in 6 Iterationen konvergiert. Koeffizienten mit einem Wert von <0,5 wurden unterdrückt.

3.4 Chancen und Risiken der Virtualisierung – Literatur- und Delphi-Studie

Die klientenbezogenen Chancen und Risiken virtueller Beratungsleistungen wurden durch Nissen et al. [20] mittels einer Kombination aus strukturierter Literaturrecherche und einer Delphi-Studie mit Teilnehmern verschiedener Branchen untersucht. Es ergaben sich die folgenden, auf der strategischen Analyseebene relevanten Ergebnisse.

Klienten verknüpfen mit virtualisierten Beratungsangeboten die Chance auf größere Flexibilität und Verfügbarkeit, höhere Arbeits- und Reaktionsgeschwindigkeit im Projekt sowie die Hoffnung, Beratungsleistungen zu günstigeren Preisen zu erhalten. Weiterhin erhoffen sich die Klienten durch die intensivere Nutzung der Digitalisierung, Ergebnisse einfacher verarbeiten und wiederverwenden zu können. Auch sieht man die Chance, besonders innovative und teilweise automatisierte Beratungslösungen in Anspruch nehmen zu können.

Dem stehen jedoch Befürchtungen und Risiken auf Klientenseite gegenüber. Diese beziehen sich einerseits auf mögliche Kommunikations-, Koordinations- und Kooperationsprobleme sowie die stärkere Abhängigkeit von technischen Aspekten bei der Zusammenarbeit und damit verbundenen Gefahren wie Datenmissbrauch und Kontrollverlust. Andererseits wird befürchtet, die Ergebnisqualität könnte leiden, da die Beziehung zwischen Beratern und Klienten sich verschlechtert, Vertrauen und Loyalität sinken und die Individualisierung der Leistungserbringung abnimmt.

Die gewonnenen Erkenntnisse fließen nun in einen mehrstufigen Vorschlag zur Analyse des Virtualisierungspotenzials von Beratungsleistungen ein. Akzeptanz auf Kundenseite wird dabei in Übereinstimmung mit Overby [23] [24] und Bruhn [5] primär als Ergebnis der erfolgreichen Umsetzung von Virtualisierungspotenzialen angesehen, ist also ein Ziel.

4 Ermittlung des Virtualisierungspotenzials von Beratungsleistungen

4.1 Überblick

Die Erfolgskette [5] virtueller Beratungsleistungen beschreibt die Stufen, die zu absolvieren sind, um Virtualisierung in der Unternehmensberatung erfolgreich umzusetzen (Abb. 1). Das Angebot und die Gestaltung virtueller Beratungsprodukte führen im Idealfall zu einer hohen Akzeptanz seitens der Kunden. In Abhängigkeit von alternativen Beratungsmöglichkeiten, empfundenen Risiken und Vorteilen, kommt es dann zu einer Nutzung des Beratungsproduktes. Während der Inanspruchnahme der Leistung beurteilt der Klient die Qualität. Entspricht diese seinen Erwartungen, entsteht Zufriedenheit. Die Zufriedenheit der Kunden stellt eine Voraussetzung dafür dar, dass auf der letzten Stufe das Beratungsunternehmen einen Erfolg mit dem angebotenen virtualisierten Beratungsleistungsportfolio realisiert.

Um eine möglichst hohe Akzeptanz der Kunden zu erzielen und deren Nutzungsabsicht zu erhöhen, ist die Gestaltung der virtualisierten Beratungsdienstleistung von großer Bedeutung. Hierzu gilt es, das Virtualisierungspotenzial zunächst in der Analyse- und Gestaltungsphase einer virtuellen Beratungsleistung fundiert zu analysieren.

Im Ergebnis erscheint es sinnvoll, die Ermittlung des Virtualisierungspotenzials im Bereich der Unternehmensberatung in drei separaten Teilschritten vorzunehmen, die nun näher dargestellt werden (Abb. 1). Die erste Stufe, die Evaluation des prozessualen Virtualisierungspotenzials, basiert auf der *Process Virtualization Theory* von Overby und untersucht, wie die Eigenschaften des Beratungsprozesses das Virtualisierungspotenzial beeinflussen. Die zweite Stufe, die unternehmensbezogene Analyse, untersucht zentrale Faktoren innerhalb des Beratungsunternehmens und wie diese auf das Virtualisierungspotenzial wirken. In der dritten und strategischen Stufe werden Chancen und Risiken der Virtualisierung aus Kunden- und Beratungssicht im Rahmen eines Business Case untersucht. Für die spätere projektbezogene Umsetzung spielen dann die Qualitätserwartungen des Kunden eine erfolgsentscheidende Rolle.

Der Evaluationsprozess startet also in Schritt 1 mit den PVT-orientierten Einflussfaktoren. Die folgenden Evaluationsschritte wenden die restlichen Kriterien an und ergänzen sie um Ergebnisse zu Chancen und Risiken der aktuellen Virtualisierungsentscheidung. Der Evaluationsprozess wird im Verlauf immer spezifischer und mündet in einen Business Case des jeweiligen Beratungshauses.

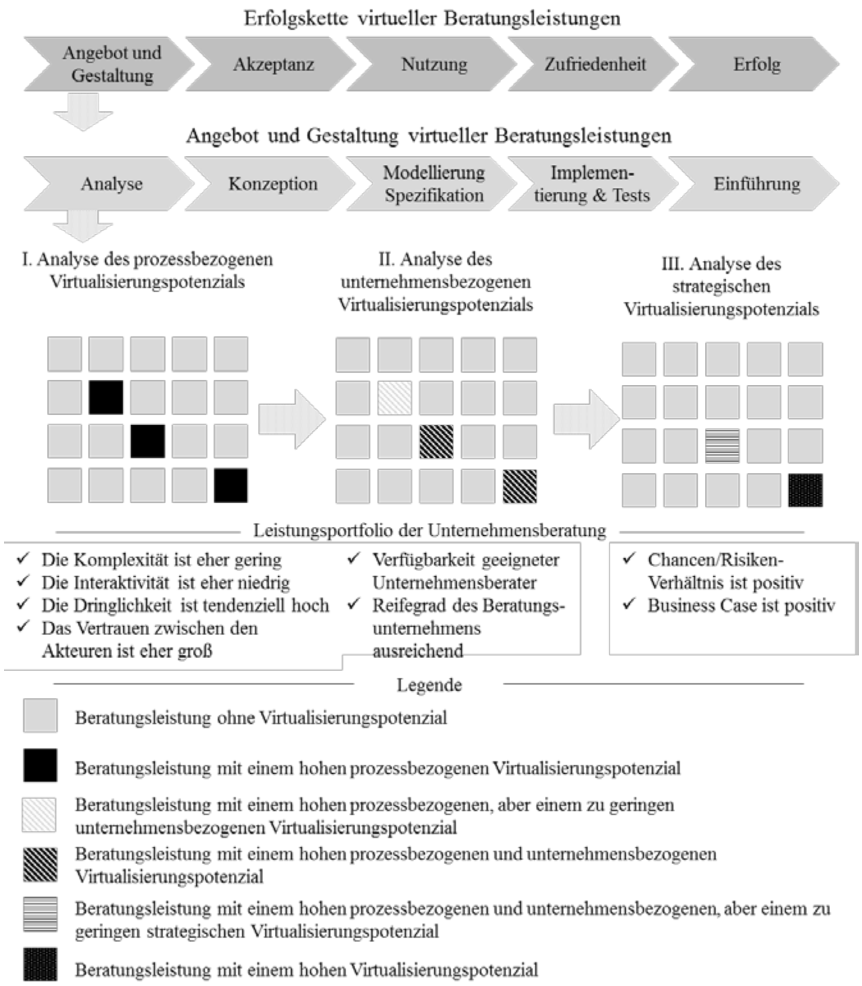


Abbildung 1. Schrittweise Analyse des Virtualisierungspotenzials von Beratungsleistungen

4.2 Analyse des prozessualen Virtualisierungspotenzials

Einfluss der Komplexität. Der erste Faktor, der für die Analyse des prozessualen Virtualisierungspotenzials entscheidend ist, ist die *Komplexität der Beratungsleistung*. Das prozessuale Virtualisierungspotenzial einer Beratungsleistung wird durch die sensorischen, beziehungsbezogenen, synchronitätsbezogenen und die kontroll-bezogenen Eigenschaften des Prozesses bestimmt. Hierauf wirkt maßgeblich die

Komplexität der Beratungsleistung. Komplexe Beratungsleistungen zeichnen sich u.a. durch eine Vielzahl von heterogenen Projektzielen, Stakeholdern, Projektstandorten und Technologien aus. Eine hohe Komplexität wirkt sich negativ auf das Virtualisierungspotenzial aus. In komplexen Beratungsszenarien sind die Anforderungen, die der Beratungsprozess an die Beziehung zwischen Berater und Klient stellt, hoch. Notwendig ist eine starke Berater-Klienten-Beziehung, die in der Lage ist, die Zusammenarbeit auch in kritischen Phasen innerhalb des Projektes zu gewährleisten. Zudem ist die Kritikalität oder das Risikopotenzial des Projektes zu beachten. Auch wenn in der Literatur bisher kein eindeutiger Zusammenhang zwischen Projekt-Komplexität und Projekt-Risiko nachgewiesen werden konnte, existiert eine verbreitete Vermutung, dass es eine Verknüpfung zwischen beiden Aspekten gibt [17]. Die Komplexität stellt zudem hohe synchronizitätsbezogene Anforderungen an den Beratungsprozess. Die Lösung komplexer Problemstellungen mit einer Vielzahl von heterogenen Informationen erfordert es häufig, zeitlich synchron miteinander zu arbeiten. Das bedeutet, Problemstellungen und Lösungen werden beispielsweise in Workshops kooperativ untersucht und generiert. Es ergeben sich zudem oft hohe sensorische Anforderungen, wenn Informationen unterschiedlichster Art verarbeitet werden müssen. Dazu zählen Sprachinformationen ebenso wie Grafiken oder Texte und insbesondere die Mimik und Gestik der verschiedenen Akteure. Komplexe Beratungsleistungen setzen auch die Kenntnis darüber voraus, wer mit wem gerade interagiert.

Einfluss der Interaktivität. Der zweite Faktor, der für die Analyse des prozessualen Virtualisierungspotenzials entscheidend ist, ist die *Interaktivität der Beratungsleistung*. Die Interaktivität der Beratungsleistung umfasst die Art, Dauer und Häufigkeit der Interaktion von Kunde und Berater sowie zwischen Beratern im Projekt. Eine hohe Interaktivität impliziert hohe sensorische und synchronizitätsbezogene Anforderungen, ebenso wie erhöhte kontroll- und identifikationsbezogene Anforderungen. Eine hohe Interaktivität schränkt somit die Virtualisierbarkeit ein. Durch die Wahl geeigneter Medien und Technologien kann dies teilweise ausgeglichen werden. Die Analyse der Antworten der Teilnehmer zeigte, dass die Interaktivität tendenziell höchstens mittelstark ausgeprägt sein sollte, damit man von einer guten Virtualisierbarkeit sprechen kann. Dieses Ergebnis ist in guter Übereinstimmung mit Erkenntnissen aus der Forschung zur Service Modularisierung. So verweisen z.B. Carlborg und Kindström [7] darauf, dass die Perspektive der Co-Kreation von Dienstleister und Kunden nicht vernachlässigt werden sollte. Starke Interaktionsanforderungen bauen auch aus Sicht der Service Modularisierung Hürden auf.

Interaktion umfasst ein breites Spektrum an Formen. Hier sollten Beratungsanbieter prüfen, welche grundsätzlichen Interaktionsformen bisher (traditionell) in einer gegebenen Beratungsleistung vorkommen: direkte persönliche Interaktion, mediale bzw. indirekte Interaktion oder automatisierte Interaktion. Häufige direkte, persönliche Interaktion zwischen Berater und Kunde deutet darauf hin, dass hier Virtualisierungspotenzial für synchrone Technologien, wie Konferenzsysteme, vorliegt [6].

Einfluss der Dringlichkeit. Die Dringlichkeit der Beratungsleistung, zu der auch die Priorität aus Kundensicht gehört, wirkt sich im konkreten Projektfall positiv auf das Virtualisierungspotenzial aus. Der Einsatz von Technologie zur örtlich unabhängigen Zusammenarbeit bringt zeitliche Vorteile mit sich, die bei dringlichen Problemstellungen tendenziell positiv wirken. Die örtliche Unabhängigkeit, die durch den Einsatz geeigneter Kollaborationswerkzeuge erreicht werden kann, fördert die schnelle und flexible Zusammenarbeit verschiedener, örtlich verteilter Projektmitglieder und ermöglicht einen schnellen Austausch von Informationen, die zur Lösung kritischer Problemstellungen benötigt werden.

Einfluss des Vertrauens. Das Vertrauen zwischen Berater und Kunde ist eines der Kernkriterien zur Beurteilung von Beratungsqualität. Die Virtualisierung von Beratungsleistungen führt zu einem veränderten Beratungsprozess, der den direkten Kontakt zwischen Berater und Kunde reduziert. Daraus resultiert ein Risiko für die Qualität der Beziehung von Beratung und Klient und insbesondere das dem Berater auf Kundenseite entgegengebrachte Vertrauen [20]. Das Vertrauen ist aus Sicht der Klienten besonders wichtig, da bei einer traditionellen wie auch einer virtuellen Beratungsleistung eine dienstleistungsspezifische Qualitäts- bzw. Informationsunsicherheit vorliegt und zum Zeitpunkt der Inanspruchnahme nur das Leistungsversprechen der Beratung existiert [20]. Großes Vertrauen der Kunden in den Beratungspartner, einerseits in das Beratungsunternehmen als Organisation und andererseits in den projektbezogenen, konkreten Berater als Person, wirkt sich positiv auf das prozessuale Virtualisierungspotenzial aus. Für die Virtualisierbarkeit von Beratungsleistungen ist es entscheidend, dass bei etablierter vertrauensvoller Zusammenarbeit zwischen Berater und Klient die beziehungs- und kontrollbezogenen Anforderungen des betroffenen Beratungsprozesses eher gering sind, da bereits eine entsprechende Berater-Klienten-Beziehung vorliegt. Die Akteure der Beratungsprozesse können daher weniger oft synchron und ohne den Bedarf häufiger Kontrollen der Zusammenarbeit miteinander interagieren. Bei stark virtualisierten Beratungsleistungen spielt vor allem das Vertrauen auf der organisationalen Ebene, also in das Beratungsunternehmen, eine Rolle. Hier kann das Reputationsmanagement der Unternehmensberatung einen wichtigen Beitrag leisten [14].

Es wird deutlich, dass die PVT alleine nicht genügt, um das Virtualisierungspotenzial praxisgerecht und nutzenstiftend bewerten zu können. Die Kriterien der PVT ermöglichen es jedoch, eine erste Einschätzung der grundsätzlichen Eignung einer Beratungsleistung für die Virtualisierung vorzunehmen. Im nächsten Schritt gilt es, dieses Urteil weiter zu schärfen und vor allem die unternehmensspezifischen Rahmenbedingungen sowie die Erwartungshaltung der Klienten zu berücksichtigen.

4.3 Analyse des unternehmensbezogenen Virtualisierungspotenzials

Einfluss des Reifegrads der Beratungsorganisation. Von großer Bedeutung für die Virtualisierbarkeit einer Beratungsleistung auf Unternehmensebene ist der Reifegrad des Beratungshauses in dieser Hinsicht. Ein hoher virtualisierungsbezogener Reifegrad eines Beratungsanbieters wirkt sich positiv auf das Virtualisierungspotenzial im konkreten Einzelfall aus. Weisen Beratungsfirmen einen hohen Reifegrad auf, dann impliziert das Erfahrungen und Know-how in der Digitalisierung von eigenen Geschäftsprozessen und Prozessen der Klienten. Dieses Wissen ist förderlich für die Entwicklung weiterer virtueller Beratungsprodukte. Firmen mit einem hohen Reifegrad können ihre Entscheidungen für oder gegen die Virtualisierung einzelner Prozesse/Prozessschritte basierend auf den schon vorliegenden eigenen Erfahrungen treffen. Das erleichtert es, das Virtualisierungspotenzial einzelner Leistungen nach Maßgabe der oben genannten Einflussfaktoren richtig zu bewerten.

Einfluss der Verfügbarkeit geeigneter Berater. Sollen virtualisierte Beratungsangebote in das eigene Leistungsportfolio integriert werden, so stellt dies Anforderungen an die Qualifikation der betroffenen Mitarbeiter. Virtuelle Beratungsleistungen erfordern neben den sozialen und fachlichen Fähigkeiten insbesondere Wissen in der Auswahl und Nutzung von passenden Informations- und Kommunikationstechnologien sowie Know-how über Formen und Besonderheiten der virtuellen Zusammenarbeit. Damit das prozessuale Virtualisierungspotenzial ausgeschöpft werden kann, müssen Berater die Wahl der passenden Medien und Werkzeuge unter Berücksichtigung der zuvor beschriebenen prozessualen Faktoren treffen. Bei der Zahl an relevanten Technologien ist hierfür fundiertes Wissen notwendig. Der Berater muss folglich in der Lage sein, die richtige Technologie für den richtigen Prozessschritt bei einem Kunden auszuwählen. Verfügt ein Beratungsanbieter über Mitarbeiter, die diesen Anforderungen gerecht werden, wirkt sich das positiv auf die Bewertung des Virtualisierungspotenzials aus. Ist das nicht der Fall, wäre es sinnvoll, zunächst in den Aufbau solcher Ressourcen zu investieren.

4.4 Analyse des strategischen Virtualisierungspotenzials

Der strategische Fit in Bezug auf das bestehende Leistungsportfolio des Beratungsanbieters ist wichtig für die Virtualisierung von Beratungsleistungen [28]. Demnach müssen Anbieter das in den bisherigen zwei Schritten identifizierte Virtualisierungspotenzial in der dritten Analysestufe auf die strategische Relevanz und Passfähigkeit prüfen. Hoher strategischer Fit wirkt sich positiv auf das Virtualisierungspotenzial aus. Die Idee einer Bestimmung des strategischen Virtualisierungspotenzials geht auf den Ansatz des strategischen Service Portfolio Managements zurück und zielt darauf ab, neue Beratungsleistungen im Hinblick auf die strategische Passgenauigkeit zum Service Portfolio des Beratungsunternehmens hin zu überprüfen und basierend darauf die Beratungsleistung und/oder das Portfolio strategisch neu auszurichten [16].

Die Forschung zur Service Modularisierung betont, dass es essentiell für Dienstleistungsanbieter ist, die Kundenwünsche zu verstehen und schon bei der Serviceentwicklung einzubeziehen [25]. Für die Entscheidung im konkreten Einzelfall ist es daher notwendig, neben den Vorteilen der Virtualisierung insbesondere deren Risiken aus Klientensicht, aber auch die Risiken der Beratung im Auge zu behalten und das Gesamturteil durch einen Business Case, der kurzfristige und langfristige Potenziale der Virtualisierung aufzeigt, weit möglichst zu objektivieren.

Die Bedeutung dieses strategischen Evaluationsschrittes wird auch durch die Analyse der Hemmnisse der Virtualisierung deutlich. So gaben 34 % der befragten

Teilnehmer der BDU-Studie an, dass der mangelnde strategische Fit ein Hemmnis der Virtualisierung in der Unternehmensberatung sei (n=493) [21].

4.5 Qualitätsanforderungen der Kunden an virtualisierte Beratung

Abschließend soll noch einmal die Perspektive des Beratungskunden betrachtet werden. Unternehmensberatungen vermarkten Leistungsversprechen in Bereichen, die für Klienten im Allgemeinen große Bedeutung haben und erhebliche Risiken mit sich bringen. Ein virtualisiertes Beratungsangebot kann langfristig nur dann erfolgreich sein, wenn die erbrachten Leistungen in ihrer Qualität die Erwartungen der Klienten erfüllen. Notwendig ist daher eine Qualitätsmessung. Dadurch kann überprüft werden, ob die Klientenerwartungen verfehlt, erfüllt oder sogar übertroffen wurden und somit die Beratungsleistung erfolgreich virtualisiert wurde oder nicht. Im letzteren Fall sollte kritisch evaluiert werden, ob die Virtualisierung generell unangemessen ist oder nur ein besserer Virtualisierungsmechanismus gefunden werden muss.

Die Untersuchung der Qualitätserwartungen von Klienten an virtualisierte Beratungsleistungen erfolgte in Nissen et al. 2015 [20] anhand einer strukturierten Literaturanalyse sowie einer anschließenden klientenbezogenen Delphi-Studie. Die von den Experten durchgeführte Evaluation und Priorisierung resultierte in einem integrierten Kriterienkatalog, der Kriterien traditioneller Beratungsleistungen und Kriterien elektronischer Dienstleistungen kombiniert. Die Delphi-Studie ergab auch, dass die Bedeutung der Qualitätskriterien für die Gesamtzufriedenheit der Kunden in Abhängigkeit des Virtualisierungsgrades variiert. So wird die Qualität einer hoch-virtualisierten Beratungsdienstleistung stärker anhand von Kriterien für die Qualität elektronischer Dienstleistungen gemessen, wohingegen eine Dienstleistung mit einem geringeren Virtualisierungsgrad stärker durch Qualitätskriterien von traditionellen Beratungsdienstleistungen bewertet wird.

5 Fazit und Ausblick

Die Akzeptanz, Nutzung und Zufriedenheit mit einer virtuellen Beratungsleistung sind gemäß Overby [23,24] die entscheidenden (ex post) Indikatoren für die Virtualisierbarkeit einer traditionellen Beratungsleistung. Damit wird aber vorausgesetzt, dass die virtuelle Beratungsleistung bereits entwickelt und im Einsatz ist. Sollen Beratungsleistungen erst noch konzipiert werden, lässt sich der zu erwartende Erfolg vorab nur schwer bestimmen. Folglich braucht es Indikatoren und ein analytisches Vorgehen zur ex ante Bestimmung der Virtualisierbarkeit einzelner Leistungen oder Prozessschritte.

Das vorgestellte dreistufige Vorgehen ist dazu ein theoretisch und empirisch fundierter Vorschlag, der sich allerdings im praktischen Einsatz erst noch bewähren muss. Hierzu sind Kooperationsprojekte mit Beratungspartnern gestartet worden, die in Kürze zu ersten konkreten Produkten führen werden. Im Ergebnis werden so Fallstudien entstehen, die den praktischen Wert der vorgestellten Methode in unterschiedlichen Beratungskontexten überprüfen und auch zeigen, inwieweit das

Vorgehen unternehmens- oder projektspezifisch konfiguriert werden muss, um den größten Nutzen zu erzielen.

Die Ergebnisse dieses Beitrags sind insofern einzuschränken, als die untersuchte Stichprobe nur den deutschen Beratungsmarkt betrachtet. Die Ergebnisse der zitierten Delphi Studien sind insoweit zu relativieren, als sie zwar den in der Literatur geforderten methodenspezifischen Umfang an Teilnehmern erreichen, dennoch aber nur das Meinungsbild einer überschaubaren Gruppe von Beratern und Klienten wiederspiegeln. Diese Ergebnisse gilt es in größer angelegten Studien in der Zukunft weiter zu vertiefen. In theoretischer Hinsicht erscheint es lohnenswert, die Themen Standardisierung und Modularisierung von Beratungsleistungen im Kontext der Virtualisierung noch einmal stärker aufzugreifen. Dabei existieren interessante Querbezüge, insbesondere zum Thema Service Modularität [7, 12, 25], wo schon verwertbare Ergebnisse vorliegen.

Virtualisierte Beratungsleistungen werden konventionelle Vor-Ort-Beratung nicht generell ersetzen. Sie sollten, dem dargestellten Entscheidungsprozess für oder gegen Virtualisierung folgend, aber als mögliche Ergänzung des Portfolios von Beratungsprodukten gesehen und fundiert analysiert werden.

Die digitale Transformation in der Beratungsbranche ist ein komplexes Unterfangen und schafft damit Barrieren gegen Wettbewerber. Wer frühzeitig relevantes Wissen akkumuliert und erfolgreiche, innovative Pilotprojekte vorweisen kann, wird langfristig voraussichtlich einen erheblichen Wettbewerbsvorteil generieren und über längere Zeit verteidigen können. Hier wird ein hohes Maß an Kreativität und strategischem Denken verlangt, das über die reine 1:1-Übertragung konventioneller Beratungsabläufe in die virtuelle Welt am Ende weit hinausreicht.

Literaturverzeichnis

1. Atteslander, P.: Methoden der empirischen Sozialforschung. 13. Aufl. Erich Schmidt Verlag, Berlin (2010)
2. Bagozzi, R. P.: The Legacy of the Technology Acceptance Model and a Proposal for a Paradigm Shift. Journal of the AIS 8, 244 – 254 (2007)
3. Balci, B.; Rosenkranz, C.: Virtual or Material, What do you prefer? A Study of Process Virtualization Theory. In: Proceedings of ECIS 2014, Tel Aviv (2014)
4. Bartlett, M. S.: The Effect of Standardization on a Chi Square Approximation in Factor Analysis. Biometrika 38, 337-344 (1951)
5. Bruhn, M.: E-Services - eine Einführung in die theoretischen und praktischen Probleme. In: Bruhn, M.(Hrsg.): Dienstleistungsmanagement. Jahrbuch 2002, 3–41. Gabler, Wiesbaden (2002)
6. Büttgen, M.: Kundenintegration in den Dienstleistungsprozess; Eine verhaltenswissenschaftliche Untersuchung. DUV, Wiesbaden (2007)
7. Carlborg, P.; Kindström, D.: Service Process Modularization and Modular Strategies. Journal of Business & Industrial Marketing 29, 313–323 (2014)
8. Churchill, G. A.: A Paradigm for Developing Better Measures of Marketing Constructs. Journal of Marketing Research, 16 (1), 64–73 (1979)
9. Cleff, T.: Deskriptive Statistik und Explorative Datenanalyse: Eine computergestützte Einführung mit Excel, SPSS und STATA. Gabler, Wiesbaden (2015)

10. Dichtl, M.: Standardisierung von Beratungsleistungen. DUV, Wiesbaden (1998)
11. Greff, T.; Werth, D.: Auf dem Weg zur digitalen Unternehmensberatung. IM+io - Magazin für Innovation, Organisation und Management, 30–34. IMC, Saarbücken (2015)
12. Dörbecker, R.; Böhmann, T.: The Concept and Effects of Service Modularity – A Literature Review. In: Proceedings 46th Hawaii International Conference on System Sciences 2013, IEEE, Piscataway/NJ, 1357 – 1366 (2013)
13. Döring, N.; Bortz, J.: Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 5. Aufl., Springer, Berlin (2015)
14. Hüttl, M.: Der gute Ruf als Erfolgsgröße. Erich Schmidt, Berlin (2005)
15. Kaiser, H.; Rice, J.: Little Jiffy, Mark 4. Educational and Psychological Measurement 34 (1), 111-117 (1974)
16. Leimeister, J.M.: Dienstleistungsengineering und -management. Springer, Berlin (2012)
17. Marle, F.: A Structured Process to Managing Complex Interactions between Project Risks. In: International Journal of Project Organisation and Management 6 (1), 4–32 (2014)
18. Nissen, V.: Consulting Research – eine Einführung. In: Nissen, V. (Hrsg.): Consulting Research, 3 – 38. Gabler Edition Wissenschaft, Wiesbaden (2007)
19. Nissen, V.; Mohe, M.; Deelmann, T.: Ziele, Anforderungen und Institutionalisierung des Forschungsfeldes Consulting Research. In: Möller, H.; Hausinger, B. (Hrsg.): Quo Vadis Beratungswissenschaft? VS-Verlag, Wiesbaden, 141 – 167 (2009)
20. Nissen, V.; Seifert, H.; Blumenstein, M.: Virtualisierung von Beratungsleistungen: Qualitätsanforderungen, Chancen und Risiken der digitalen Transformation in der Unternehmensberatung aus der Klientenperspektive. In: Deelmann, T.;Ockel, D.M. (Hrsg.) Handbuch der Unternehmensberatung. 25. Erg.-Lfg.. Erich Schmidt Verlag, Berlin (2015)
21. Nissen, V.; Seifert, H.: Virtualisierung in der Unternehmensberatung – eine Studie im deutschen Beratungsmarkt in Kooperation mit dem Bundesverband Deutscher Unternehmensberater BDU e. V., Bonn (2016)
22. Nunnally, J. C.; Bernstein, I. H.: Psychometric Theory. 3. A., New York (1994)
23. Overby, E.: Process Virtualization Theory and the Impact of Information Technology. Organization Science 19 (2), 277–291 (2008)
24. Overby, E.: Migrating Processes from Physical to Virtual Environments. In: Dwivedi, Y. K. et al. (eds.): Information Systems Theory, Vol. 1. Springer, New York, 107–124 (2012)
25. Rahikka, E.; Ulkuniemi, P.; Pekkarinen, S.: Developing the Value Perception of the Business Customer through Service Modularity. Journal of Business & Industrial Marketing 26, 357–367 (2011)
26. Seifert, H.; Nissen, V.: Virtualisierung von Beratungsleistungen: Stand der Forschung zur digitalen Transformation in der Unternehmensberatung und weiterer Forschungsbedarf. In: Nissen,V.; Stelzer, D.; Straßburger, S.; Fischer, D. (Hrsg.): Proc. MKWI2016, 1031-1040. ilmedia, Ilmenau (2016)
27. Webster, J.; Watson, R.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26 (2), 13-23 (2002)
28. Wurdack, A.: E-Consulting - Entwicklung eines Rahmenkonzeptes; Aufbau und Darstellung einer E-Consulting-Lösung im Beratungsunternehmen der Zukunft. Dissertation, Universität Mannheim (2001)

Navigating Digital Innovation – The Complementary Effect of Organizational and Knowledge Recombination

Sebastian Dürr[1,2], Heinz-Theo Wagner[1], Tim Weitzel[2], and Daniel Beimborn[3]

[1] German Graduate School of Management & Law, Management and Innovation, Heilbronn, Germany
{sebastian.duerr,heinz-theo.wagner}@ggs.de
[2] University of Bamberg, Chair of Information Systems and Services, Bamberg, Germany
{sebastian.duerr,tim.weitzel}@uni-bamberg.de
[3] Frankfurt School of Finance & Management, Management Department, Frankfurt, Germany
{d.beimborn}@fs.de

Abstract. This paper reports findings from an exploratory series of case studies undertaken to better understand the impacts of digitization on organizational and knowledge recombination. While 'digitalized firms' are expected to frequently update their knowledge (e.g., big data analyses, 'smarter' products) and organizational structures (e.g., agile team structures, open innovation approaches), we know little about the interrelations between recombining knowledge and organizational capabilities. Therefore, we collected data from 19 interviewees in 8 different firms from four industries. Our study integrates theoretical notions from the literature on organizational and knowledge recombination theory and categorizes the emerging shifts arising from digitization. Our results suggest that the identified changes in the digital age manifest as increased proximity to the customer (by fully digitizing the customer interface and digital co-creation) and celerity to the markets (cross-organizational teams, and collaborations with start-ups and competitors) affecting the whole organization.

Keywords: digital innovation, case studies, recombination, celerity, proximity.

1 Introduction

Ever since the seminal writings of Schumpeter, recombination is at the heart of innovation practices [1]. Schumpeter noted that innovation "consists to a substantial extent of a recombination of conceptual and physical materials that were previously in existence" [2, p. 88]. Accordingly, recent research commentaries on innovation in the digital age (e.g. [3]) also portray digital innovation in terms of recombination. Yoo et al. describe digital innovation as "the carrying out of new combinations of digital and physical components to produce novel products" [4, p. 725].

It is widely acknowledged that for a firm to 'go digital' will involve updating knowledge (e.g., big data analyses, 'smarter' products) and altering structures (e.g., agile team structures, open innovation approaches). Still, there is a need to better

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Dürr, S.; Wagner, H.-T.; Weitzel, T.; Beimborn, D. (2017): Navigating Digital Innovation - The Complementary Effect of Organizational and Knowledge Recombination, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1363-1377

understand *what* components need to be recombined and to what extent there are interrelations between recombining knowledge and organizational capabilities. The innovation literature often scrutinizes the ability of firms to recombine resources in general [1], knowledge resources [5], IT resources with other resources [6, 7], or digital and physical components [8]. For example, Carnabuci and Operti [9] investigate recombinant capabilities of firms that allow them to recombine existing technologies for innovations.

However, the extant literature in the field of recombination and innovation mainly focuses on knowledge and component recombination, neglecting potential complementarities to the organizational perspective. Although Galunic and Eisenhardt [10] mention that delineations of competency areas may affect resource combinations, and Karim and Kaul [11] examine how the recombination of business units within a firm affects firm innovation, we know little about the interplay of recombining knowledge and organizational elements. We hence aim at identifying shifts in digitizing firms and disclosing how those firms navigate their knowledge and organizational capabilities from traditional waters into a digital environment by knowledge and organizational recombination. We are particularly interested in how this recombination affects innovation success, i.e. the commercialization of new knowledge in form of new products or services [12]. Following the recent call of Henfridsson et al. [13] that "it would, therefore, be useful to examine […] the relationship between organizational design and product design" [p. 39] our research question is: *How does the interplay of organizational and knowledge recombination relate to innovation success within a digital context?*

Methodologically, we conducted exploratory case studies with 19 interviews in 8 organizations to derive categories for organizational and knowledge recombination and identify complementarities among them. Theoretically, we build on Carnabuci and Operti's [9] distinction of recombinant reuse (recombination of familiar knowledge) and recombinant creation (recombination of unfamiliar knowledge). Concerning organizational recombination, we build on Karim and Kaul [11] but go beyond their notion of structural recombination by also considering inter-organizational relationships as identified by Mintzberg [14]. Together, this allows us to incorporate the largely neglected complementarities between the knowledge and organizational based perspectives in innovation research that, as the results show, shape the modern digital firm.

The next sections review the literature on recombination theory from a knowledge and organizational perspective, introduce and analyze the case studies and then discuss the main findings, their limitations, and implications for future research.

2 Recombination - A Theoretical Background

2.1 Recombinant Reuse and Recombinant Creation

Carnabuci and Operti [9] provide a granular perspective on recombination by distinguishing between recombinant creation and recombinant reuse. *Recombinant creation* is the "ability to envision and create combinations using technologies that […]

have never [been] combined before" [6, p. 1592]. An example is the "fitness tracking" services enabled by connecting Nike's running shoes and Apple's computing products, two technologies that the firms had never previously integrated [15].

Recombinant reuse is the "ability to refine and reuse systematically known technological combinations to solve new problems and develop new applications" [6, p. 1592]. Henderson and Clark [16] call this architectural innovation as the recombination is used to replace existing components because of an inferior cost structure and/or performance issues. An example of recombinant reuse is Apple realizing shortly after the invention of the iPhone that combining its operating system and its mechanical components, e.g. its camera, offered opportunities to develop a wide range of novel applications [17]. Thus, they progressively deepened their understanding of this technological combination and generated a string of advanced technological devices, including gyroscopes and accelerometers.

2.2 Organizational and Knowledge Recombination

Henri Poincaré pointed out that new knowledge stems from existing knowledge, and is (or builds on) the combination of existing knowledge [18]. Still, controversies remain on where the knowledge elements should be selected from. Some researchers (e.g. [19]) stress that industries are clearly divided and that various technical fields are strongly separate so that knowledge inside a field can hardly be understood by outsiders. Therefore, new knowledge is often built on existing knowledge elements within such fields [20]. In contrast, Hargadon [21] suggests that technical convergence through combining knowledge from different fields can create huge market returns. The idea is that combining more diverse knowledge sources challenges actors and existing concepts and thus stimulates innovation [22]. This notion is elaborated for the digital era by Lyytinen et al. [23] and Nambisan [3]. They emphasize that convergence and recombination of knowledge are particularly advantageous for digital innovations as the malleability of such technology [13] allows for greater flexibility in the alteration of digital products from a knowledge perspective.

Nevertheless, the recombination of knowledge from various fields may be obstructed by inter- and intra-organizational boundaries between firms, business units, or even departments [24]. Hence, Karim and Kaul [11] examine the effect of the recombination of organizational structural elements ("structural recombination") on innovation from an organizational perspective. They find that structural recombination complements knowledge recombination by dissolving unit boundaries, thereby enabling intra-organizational knowledge synergies. Similarly, previous work has shown that the transfer and sharing of knowledge across internal boundaries (i.e. amongst different units within an organization) is an important source of firm innovation and competitive advantage [1]. The essence of this strand of research is that a firm that wants to realize intra-organizational knowledge synergies beyond the explicit transfer of knowledge between units should alter its structure, dissolve internal boundaries and reshuffle activities among units [25]. Karim and Kaul refer to these changes as structural recombination, defined as "changes in business units as their resources and market activities are reorganized by merging units together, generally through the absorption of one unit into another unit or the formation of a new business unit by combining existing units" [19, p. 441]. This highlights that structural recombination is a subset of the larger set of organizational reconfigurations that have been studied in the prior literature and that include the addition and deletion of units as well as their recombination [26].

3 Methodology

We followed an exploratory case study approach to understand how the interplay of organizational and knowledge recombination is related to innovation success. We chose case studies as they are particularly suitable to answer "how" and "why" research questions and when the relationship between context and phenomenon is unknown [27][28]. For data collection, we conducted 19 interviews following semi-structured guidelines with open-ended questions to assure openness for every possible research direction. We contacted senior managers responsible for strategy, R&D, innovation, IT, or marketing/sales from various firms that may be expected to be key informants in the areas of interest. Usually, we collected data in a firm if three different senior managers agreed to participate. In case that one manager was able to cover the entire area of interest because of his background, we rested with one interviewee. The organizations were not limited to any industry in order to learn about differences in several sectors. Table 1 provides an overview of the cases, industries, and interviewee position:

Table 1. Case, Industry, ID, Interviewee Position and Length

Case	Industry	ID	Interviewee Position	Length [min]
A	Manufacturing	IP01	Innovation Manager	59
B	Financial Services	IP02	Innovation Manager	68
		IP03	Head of Product Management	66
		IP04	Chief Technical Officer	57
C	Publishing	IP05	Head of Business Development	72
		IP06	Deputy General Manager	70
		IP07	Chief Technical Officer	66
D	Financial Services	IP08	Marketing Manager	72
		IP09	Head of IT & Organization	83
		IP10	Chief Executive Officer	45
E	Manufacturing	IP11	Product Group Manager	91
		IP12	Head of Automation and Controls	121
		IP13	Director of Technology Management	54
F	Financial Services	IP14	Chief Digital Officer	57
G	Financial Services	IP15	Chief Executive Officer	65
H	Manufacturing	IP16	Chief Digital Marketing Management	53
		IP17	Chief Financial Officer	46
		IP18	Chief Marketing Manager	53
		IP19	Chief Executive Officer	45

We carried out the interviews mostly onsite involving two or three of the authors as interviewers. Our interview guideline is designed as follows: First, we asked about any recent digital initiative. We particularly probed into specificities of the company's innovation process, critical success factors, involved units, and the corresponding team structure. Furthermore, we gathered information on the wider organizational structure of the company, its governance, and leadership structure. The interviews took place from November 2015 to July 2016. All interviews were transcribed, and project documentations, related reports, company's financial statements, off-record notes, and observations were used to augment and triangulate the interview data. Data analysis proceeded through reiterations between looking for meanings in the data, writing descriptions, coding and revisiting literature. Following Miles and Huberman's [29] recommendations, this data analysis process was facilitated through the building of data displays in the form of tables and matrices (through coding in MaxQDA v.12.1.4) to refine the concepts identified, and the development of tentative conclusions to depict the emerging shifts.

The data analysis began with descriptive codes (or open coding) as soon as the first interviews were transcribed, and was done inductively, seeking to reflect the data as closely as possible. This stage led to the identification of over 280 descriptive coded statements. These statements were then organized into 27 different categories, such as firm specifics, initiatives due to digitization, agility, digital innovation, structure of an organization, organizational recombination, strategy etc. (these being the early interpretative codes). At this point, the focus was on interpreting the data to search for relationships and patterns and facilitate the next stage, pattern coding. The process was highly iterative, moving between data, interpretation, and theory. As our output, we identified two central emerging shifts that we structured in *knowledge recombination* and *organizational recombination.* Focusing on these shifts, we identified complementarities among them as we looked back at the data which are presented in the next section.

4 Results

4.1 Knowledge Recombination in the Digital Age

Collecting Data to Gain Knowledge about Customers in the Digital Age. In the conversations with the firms, we noted their necessity to collect increasing amounts of customer data to better understand customer needs and wants and to eventually adjust existing or create new services. Likewise, the CTO of a large online bank (IP04) remarks: "We want to gain more information about our customers […]. We'd like to gain knowledge about their living conditions and their financial situation […]. By today, we do not have the comprehensive view of the customer […] that we would like."

For this endeavor, the firm intends to track all digital activity of their customers through their connected digital products or digital services (e.g. usage data, transaction data, location data) and to combine it with their corresponding customer master data (e.g. profession, age). Another new challenge, according to the interviewees, is to complement their conventional user data (e.g. transaction data, product usage data) with secondary data that might initially seem unrelated (IP04: “We would even add further information such as weather data”) [30] but offer greater possibilities in predicting or concluding certain developments or desires of a customer [31]. To summarize, companies use new data sources to extend the knowledge on their customers for deeper analysis and to recombine this knowledge.

Knowledge Recombination via Recombinant Creation. Collecting primary data through tracking the digital activity of customers can lead to the development of new products or services as the behavior of customers may provide the firm with insights unavailable before [30–32]. Accordingly, the innovation manager (IP02) of a bank stated: “We collect all this data because it is an effective way to delineate innovations by identifying what is missing [for our customers] and where they have problems.”

This manager explained that his company analyzes when customers are aborting certain digital processes, and try to enhance these digital processes by providing their customers with, for instance, improved web page structures or more options to choose from. Then, such data is complemented by external data to further augment a firm’s knowledge about their customers and to create a comprehensive understanding of customer needs and wants. This, ideally, makes it possible to provide fitting offerings without requiring customers to enter lots of data: “In the future, [the customers] do not need to do anything because we will give [them] recommendations” (IP04).

Such recommendations can be innovative solutions that are being tailored based on the previously tracked and analyzed the digital activity of a customer combining knowledge from conventional (master data, digital behavior) and secondary (environmental) data. This way, firms add value by digitally engaging with their customers to get ideas for novel products, ideally, as we learned, in an automated way. It is of high priority to satisfy the customer’s needs following the core marketing idea of adjusting value propositions to consumer needs [33], and now this appears to be even more relevant as customer loyalty seems to become less important (IP14): “it does not matter if we want to [do this] or not. The customers will base their [banking choice] decision on where the banking experience is the easiest, most convenient or the greatest.” These insights indicate that companies try new combinations in products or services (recombinant creation) through recombining knowledge which is based on collecting and analyzing data through tracking the digital activity of customers (knowledge recombination).

Knowledge Recombination via Recombinant Reuse. Still, tracing digital activity is not only about becoming a “comprehensive advisor” but also about the co-creation with customers [33, 34]. This assumes that different entities (e.g. firms and consumers) cooperatively integrate their operant resources (e.g. knowledge, skills, and technology) [3] within a joint process to generate value [35]. A marketing specialist (IP11) introduces his organization’s practices as operated in the agricultural sector when they equip farms: “We do not develop one [stand-alone milking] robot per customer […]

[where the milking robot is just one component of the entire farm]. Yet, the information relevant to our sales process are how does the customer want his farm? What is his daily need? […] Therefore, we develop his farm in co-creation with the customer to ensure that he wants to have it."

He underpins that the company's approach to solving new problems (i.e. integrating a stand-alone milking robot) is using known combinations, i.e. through value co-creation with the respective customer. Therefore, the firm involves such customers as closely as possible to reduce technological uncertainty [36]: "We do not always know what we want to develop. Thus, we research a lot with our customers. We would run an algorithm in test mode to observe what this means to the customer. This way, we could improve a lot of functions, […] and increase performance" (IP11).

Hitherto, co-creation is not a novel concept [34]. But with digitization and being permanently connected to the customer it allows for co-creation on a large scale, as pinpointed in this statement (IP03): "For testing user satisfaction, we got our 'advisory customers'. This is an insider community of some five hundred customers which we survey continuously through digital channels."

From this testing, the company derives how customers use new products, and identify spaces for improvement. In the following, its CTO (IP04) explains how this improvement occurred in an agile way by recapturing the development of a successful app:

"We would try to only offer this core feature […]. This way, we can first test how the product is being perceived […] By now, I think, that the third version of the app is available – within only four or five months […]. Previously, we probably would have released the first version only late in summer, realizing all the other features that we have thought of. In this example, we have realized the very first functional version, and […] rereleased further functions."

Until recently, the traditional testing of products, i.e. imitating how clients will use products, was the means to identify potential weaknesses [37] and to ensure that new products conform with their intended functionalities [31]. In the digital age, however, companies monitor and improve their products continuously, i.e. they might compile real performance data from all products which may provide them with insights on design problems that the artificial testing did not reveal.

Accordingly, we observe that companies solve the problems of their customers (recombinant reuse) in digital co-creation through the recombination of known technological combinations supported by the collection and analysis of customers' digital activity (knowledge recombination). Further, we derive that these activities allow for greater proximity of a firm to every individual customer through digital technology in an unprecedented automatized way. *Proximity* connotes thoroughly analyzing the digital behavior of customers as a resource for innovation ideas, as a co-creator in the development and design of products and services, and as a user in testing the product or in helping other users to learn about the product or service [3]. This means that companies focus on creating market offerings after having digitally analyzed their customers' behavior (e.g. product usage via tracking digital activity [30, 31, 38]) and before tailoring specific affordances through a recombination of the gained customer knowledge (recombinant creation).

4.2 Organizational Recombination as the Complement of effective Knowledge recombination in the Digital Age

Restructuring the Organization in the Digital Age. In our eight case studies, we were sensitized that identifying new knowledge combinations through digital analysis, and solving problems of customers through digitally gathered knowledge, are associated with organizational recombination and eventually aimed at increased organizational agility. For instance, the agile development of products and services requires new units or functions [31, 39]. Likewise, the Chief Digital Officer of a retail bank (IP14) confers the sheer importance of becoming "more agile" and highlights how his firm's senior management team infuses his firm: "We have identified [digital] strategic fields which we elaborate on cross-divisionally, i.e. with teams created from the different areas of the company, with the aim to develop or identify a greater need for change."

By becoming more agile, firms want to be able to sense and respond quicker to customer-based opportunities [40, 41] to timely incorporate ('recombine') the knowledge gained from their digital activities. Besides, the CTO (IP07) of a large weekly explains another advantage of agile development approaches: "You just get more insight, you get insights faster, also when something goes out of control".

Following this prospect, firms may become more agile in various areas, such as customer-based processes or interactions with supply chain partners, or to increase competitiveness [41].

Organizational Recombination via Recombinant Creation. Firms have long tried new organizational combinations with external partners in various ways [11, 26]. Our case studies confirmed this practice, yet the objective slightly deviates. Hence, companies would continue with the traditional integration of external partners by acquisition [11] ("Fintechs[1] are integrated with quite banal cooperation models. We honestly cooperate with Fintechs for many years" (IP14)), but at the same time they are trying to adopt the agile spirit of start-ups: "It is of interest for us to experience entrepreneurship. The spirit, the gut decisions, this enthusiasm […]. I believe this 'spirit transfer' in the sense of digital transformation is extremely important. Especially for larger companies."

This spirit appears to complement the co-creation between a firm and its customers (knowledge recombination) and may strengthen through solutions integrated through collaborations with startups or competitors (organizational recombination). The rationale is that startups or competitors may provide novel, legacy-free solutions for the customers of a "traditional" firm that the incumbent may have difficulties to develop due to a lack of skills, its strong hierarchy, or slow decision-making processes [14]. Hence, all firms reported integrating startups, while a subset creates incubator structures (in cases A, B, F, and G) to attract and support startups at their very founding stage [42]. Firm A also launches small spin-offs to save resources (e.g. cost, and manpower) as internal efforts are being decreased, and allow for more agile realizations compared to those of his organization due to "less organizational legacy". However, a

[1] Fintech is a synonym for start-ups in the financial industry.

product group manager (IP11) also introduced us to their shift towards operating on platforms on which they collaborate with their competitors, aiming at providing the customer with solutions to their individual problems regardless where the solution is coming from:

"We see farmPage[2] as a platform. We would appreciate if our competitors would be on it, too […]. The more customers, […] the more valuable is what you can analyze from it." With these words, the manager highlights the notion of collaboration with competitors. He further notes that in the digital age customers are less involved with minding the origin of a solution to their problems, rather than having an organizational combination that provides them with a solution. Besides, the last part of the quote underlines how the collaboration with customers and competitors may serve as a new source for knowledge recombination. Hence, we note that companies form various new organizational combinations (recombinant creation) through the integration of external partners, i.e. start-ups, incubators, spin-offs or competitors, to increase the celerity in developing innovations.

Organizational Recombination via Recombinant Reuse. Besides forming new organizational combinations with external digital partners, the interviewed firms alter their internal structures by creating cross-functional teams. This brings knowledge resources closer together that were previously separated, and removes internal boundaries between them thus increasing the likelihood of knowledge recombination [43]. One manager (IP02) sketches how such teams are constituted in digital initiatives: "We do not have these silos anymore, which we had […]. Now, all parties are in the product development: marketing, IT, legal and product managers." This structural recombinant reuse (forming cross-organizational teams) within the organizational structure (organizational recombination) fosters the thorough engagement with a firm's customers as the analysis of customer data and environmental data can be facilitated by various experts from heterogeneous fields (knowledge recombination) allowing a new view on existing products) [9].

In addition, the companies restructured their physical workspace to further support this cross-functional collaboration. With wide spaces without boundaries and people mingling together, knowledge is more easily shared within units than between them [44], even when activities within a unit are distant from each other [45]. One manager illustrated this for his firm (IP03): "We have started to completely restructure many of our workspaces. […] We created larger spaces because many small offices existed in which employees got lost […] Now, the people focus on one topic, and the cycles of product development shrunk significantly […]. In addition, the workspace development is very, very good for our staff […] captured on the balance sheet." Overall, fewer divisional boundaries enable greater and wider exploration of knowledge through richer communication and joint development [46] by forming cross-departmental teams, and by co-locating people.

In the digital era, another emerging shift comes from the formation of new positions, and dedicated managers to consolidate the digital activities of a firm [30, 31]. Hence,

[2] Software product name changed to assure anonymity of all interview partners and related firms.

several managers (IP02, IP08, IP04, IP14, IP16) told us they were hired to introduce initiatives to make their company more "digital". Other companies (cases E and F) even instantiate new digital structures, such as a digital office equipped with the competence to make their company more agile. Correspondingly, a newly appointed chief digital officer (IP14) describes his current position: "Our Chief Digital Office reports directly to the CEO. I am a member of a six-member management team in the CDO department [...]. Since May last year, we made an organizational change, [...] so that I get the space to deal with the major disruptive changes in our business [...]. Now I deal with the digitization [...]. In particular, the management of innovation and in this context the exchange with startups, but also with other potential partners in the digital context."

He summarizes how the integration of external partners is fortified through the firm's internal restructuring. This is complemented by digitizing the customer interface for digital customer analyzes (aimed at knowledge recombination) because it simplifies the integration of external partners (organizational recombination). Through the digitization of an interface, a firm defines certain standards (i.e., APIs) which also can be used when incorporating startups, respectively their solutions into a company.

Overall, we note that companies recombine their internal organizational structure, responsibilities, and functions (organizational recombination) to be quicker at solving problems in the digital age (recombinant reuse) and to form new organizational combinations with external partners (recombinant creation). Digital co-creation (knowledge recombination) may even support this by spurring quicker product or service releases for the customers if the organization uses agile methodologies in heterogeneous teams (organizational recombination), which is reportedly integrated into the daily business practices by all managers. Our findings are depicted in Figure 1:

Recombinant Type	Area of Recombination		
	Knowledge Recombination	Organizational Recombination	
Recombinant Creation	[a] Digitizing the customer interface (**Smart Customer Analysis**) Reported in: A, B, E, F	[b] Becoming more open to external partners **(Startups, Competitors)** Reported in: A, B, C, D, E, F	
Recombinant Reuse	[c] Solving problems through digitally gained knowledge (**Digital Co-Creation**) Reported in: B, E	[d] Becoming more agile in development (**Cross-Organizational Teams**) Reported in: B, C, D, F, H	
Identified Complementarities	Smart customer analysis [a] enhances in cross-organizational teams [d]		IP02
	Digital Co-Creation [c] improves through integrating solutions of Startups or Competitors [b]		IP11
	Digitizing the customer interface [a] simplifies the integration of Startups or Competitors [b]		IP08
	Digital Co-Creation [c] is quicker realized in agile development [d]		IP06

Figure 1. The Emerging Shifts in the Digital Age and their Complementarities

5. Discussion

We theoretically and empirically analyzed complementary organizational and knowledge capability recombination as pillars of a firm's digitization initiative. The results reveal consistent shifts in organizational and knowledge recombination.

Organizations recombine knowledge elements based on customer analysis identified in an unprecedented way which is enabled through fully digitizing the customer interface, to achieve greater proximity for deriving innovations, and to co-create products which are quickly realized. As product co-creators, customers often develop and refine product and service features that the firm incorporates into future versions [47]. Finally, customers who become expert users often discover new ways to use the product as well as shortcuts and other methods to enhance the overall value of the product.

Organizations recombine organizational elements by forming cross-functional teams, and by collaborating with external partners to achieve greater celerity in developing new products reflecting the customer preferences identified with the help of digitalized customer interfaces. Celerity echoes the imperatives of fast, recurrent product design, continuous product operation, and ongoing product upgrades [39] that create a need for new functional groups. In the software industry, collaborative and cross-functional software development methods like scrum have long been established. They bring together software engineers and R&D staff with IT, manufacturing, and service staff who deal with operational matters [31]. This approach rests on teams that manage product updates, shorten product cycles, and deliver new services and enhancements after the sale.

In this manuscript, we theoretically and empirically analyzed the complementarities of organizational and knowledge capability recombination. Thus, we contribute to existing research in organizational and knowledge recombination in two ways: First, we highlight that not only intra-organizational recombination of organizational elements but also inter-organizational recombination plays a role in the digital age. In addition, we extend this notion by exploring that not only investigating unit boundaries is necessary going beyond the work of Karim and Kaul [11]. Second, looking for key shifts for firms in the digital age we identify celerity and proximity as pillars of a firm's digitization initiative. Further, the results highlight complementarities between knowledge and organizational recombination for innovation that have only attracted scant recognition in the literature so far [13]. Our cases indicate that these complements may exemplify ways of becoming faster and more agile when developing digital innovations.

The results suggest some interesting implications and avenues for future research on the recombination of complementary knowledge and organizational elements (cf. Figure 1). For knowledge elements, the number of accessible knowledge elements constitutes a knowledge recombination space. Hence, combination means building networks of knowledge elements, and recombination means creating new networks. Analogously, a firm's organizational elements (i.e. structural elements such as units, teams, and cross-functional teams, both inter-organizational and intra-organizational, power distribution, and decision-making modes [14]) constitute the organizational recombination space. Recombination then means managing boundary conditions and interventions that influence the creation of new knowledge networks [11, 48]. Drawing on the knowledge management and enterprise governance and change literature, the findings could offer a first framework for more theoretical and empirical work on digitization.

For practice, the results imply that firms increasingly pursue agile methodologies to spur faster results and use digital channels to either more thoroughly analyze or engage with a firm's customers, competitors or partners. These practices may provide direction for practitioners when navigating the enterprise into digital waters.

Future research may increasingly aim at each dimension of the emerging shifts in the digital age as illustrated in Figure 1. Hence, further organizational elements such as an organization's culture or power structure may shift in a digital context. It will also be necessary to transcend our focus on European companies and include firms from other continents and cultures.

Our research needs to be reflected in the light of its limitations. We did not use a chronological order nor an order based on popularity due to the heterogeneity of our interview partners. Instead, we ordered the findings by forming a theoretical lens (cf. Figure 1), and by following this lens' aspects to classify the identified shifts. Second, there may be a bias as two, respectively three, people attending the interviews may have intimidated the interview partners. However, we avert this by selecting one main interviewer to not overcharge the interviewee.

6. Conclusion

Our research question was: "*How does the interplay of organizational and knowledge recombination relate to innovation success within a digital context?*" The analysis mainly shows that organizational and knowledge recombination complement each other leading to increased proximity to a firm's customers and to increased celerity to markets to develop novel innovations when navigating in an increasingly digital environment.

References

1. Galunic, C., Rodan, S.: Resource recombinations in the firm: knowledge structures and the potential for Schumpeterian innovation. Strateg. Manag. J. 19, 1193–1201 (1998).
2. Schumpeter, J.A.: The theory of economic development: an inquiry into profits, capital, credit, interest, and the business cycle. Harvard Econ. Stud. 46, xii, 255 (1934).
3. Nambisan, S.: Information Technology and Product/Service Innovation. J. Assoc. Inf. Syst. 14, 215–226 (2013).
4. Yoo, Y., Boland, R.J., Lyytinen, K., Majchrzak, A.: Organizing for Innovation in the Digitized World Organizing for Innovation in the Digitized World. Organ. Sci. 23, 1398–1408 (2012).
5. Kogut, B., Zander, U.: Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. Organ. Sci. 3, 383–397 (1992).
6. Jarvenpaa, S.L., Leidner, D.E.: An Information Company in Mexico: Extending the Resource-Based View of the Firm to a Developing Country Context. Inf. Syst. Res. 9, 342–361 (1998).
7. Mitchell, V.L.: Knowledge Integration and Information Technology Project Performance. MIS Q. 30, 919–939 (2006).
8. Yoo, Y., Henfridsson, O., Lyytinen, K.: The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research., 10.1287/isre.1100.0322, (2010).
9. Carnabuci, G., Operti, E.: Where do firms' recombinant capabilities come from? Strateg. Manag. J. 34, 1591–1613 (2013).
10. Galunic, D.C., Eisenhardt, K.M.: Architectural Innovation and Modular Corporate Forms. Acad. Manag. J. 44, 1229–1249 (2001).
11. Karim, S., Kaul, A.: Structural Recombination and Innovation: Unlocking Intraorganizational Knowledge Synergy through Structural Change. Organ. Sci. 26, 439–455 (2015).
12. Joshi, K.D., Chi, L., Datta, A., Han, S.: Changing the Competitive Landscape: Continuous Innovation Through IT-Enabled Knowledge Capabilities. Inf. Syst. Res. 21, 472–495 (2010).
13. Henfridsson, O., Mathiassen, L., Svahn, F.: Managing technological change in the digital age. J. Inf. Technol. 29, 27–43 (2014).
14. Mintzberg, H.: The Structuring of Organizations. Prentice-Hill, Englewood Cliffs, NJ, USA (1979).
15. Yoo, Y.: Computing in Everyday Life: A Call for Research on Experiential Computing. MIS Q. 34, 213–231 (2010).
16. Henderson, R.M., Clark, K.B.: Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. Adm. Sci. Q. 35, 9–30 (1990).
17. Tilson, D., Lyytinen, K., Sørensen, C.: Research Commentary—Digital Infrastructures: The Missing IS Research Agenda. Inf. Syst. Res. 21, 748–759 (2010).
18. Carnabuci, G., Bruggeman, J.: Knowledge Specialization, Knowledge Brokerage and the Uneven Growth of Technology Domains. Soc. Forces. 88, 607–641 (2009).
19. Dosi, G.: Technological paradigms and technological trajectories. A suggested

interpretation of the determinants and directions of technical change. Res. Policy. 11, 147–162 (1982).
20. Henderson, R., Jaffe, A., Trajtenberg, M., Thompson, P., Fox-Kean, M.: Patent Citations and the Geography of Knowledge Spillovers. Am. Econ. Rev. 95, 461–466 (2005).
21. Hargadon, A.B.: Firms as knowledge brokers : Lessons in pursuing continuous innovation. Calif. Manage. Rev. 40, 209–227 (1998).
22. Miller, D.J., Fern, M.J., Cardinal, L.B.: The use of knowledge for technological innovation within diversified firms. Acad. Manag. J. 50, 308–326 (2007).
23. Lyytinen, K., Yoo, Y., Boland, R.J.: Digital product innovation within four classes of innovation networks. Inf. Syst. J. 26, 47–75 (2016).
24. Tsai, W.: Social Structure of "Coopetition" Within a Multiunit Organization: Coordination, Competition, and Intraorganizational Knowledge Sharing. Organ. Sci. 13, 179–190 (2002).
25. Brown, S.L., Eisenhardt, K.M.: The Art of Continuous Change: Linking Complexity Theory and Time-paced Evolution in Relentlessly Shifting Organizations. Adm. Sci. Q. 42, 1–34 (1997).
26. Karim, S.: Modularity in organizational structure: The reconfiguration of internally developed and acquired business units. Strateg. Manag. J. 27, 799–823 (2006).
27. Eisenhardt, K.M.: Building Theories from Case Study Research. Acad. Manag. Rev. 14, 532–550 (1989).
28. Yin, R.K.: Case Study Research: Design and Methods. Appl. Soc. Res. Methods Ser. 5, 219 (2009).
29. Miles, M., Huberman, M.: Qualitative data analysis: An expanded sourcebook. Sage Publications, Thousand Oaks, CA (1994).
30. Westerman, G., Bonnet, D., McAfee, A.P.: Leading Digital. Havard Business Review Press, Boston, MA (2014).
31. Porter, M.E., Heppelmann, J.E.: How Smart, Connected Products are Transforming Companies. Harv. Bus. Rev. October 20, 96–114 (2015).
32. Bennis, W.: Leadership In A Digital World: Embracing Transparency And Adaptive Capacity. MIS Q. 37, 635–636 (2013).
33. Shah, S.K.: Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development. Manage. Sci. 52, 1000–1014 (2006).
34. Boulding, W., Staelin, R., Ehret, M., Johnston, W.J.: A Customer Relationship Management Roadmap : What Is Known , Potential Pitfalls and Where to Go. J. Mark. 69, 155–166 (2005).
35. Vargo, S.L., Lusch, R.F.: Service-dominant logic: Continuing the evolution. J. Acad. Mark. Sci. 36, 1–10 (2008).
36. Fleming, L.: Recombinant Uncertainty in Technological Search. Manage. Sci. 47, 117–132 (2001).
37. Bailey, D.E., Leonardi, P.M., Barley, S.R.: The lure of the virtual. Organ. Sci. 23, 1485–1504 (2012).
38. Brynjolfsson, E., McAfee, A.P.: The Second Machine Age - Work, progress, and prosperity in a Time of Brilliant Technologies. WW Norton & Company (2014).
39. Fitzgerald, B., Stol, K.J.: Continuous software engineering: A roadmap and agenda. J. Syst. Softw. 21, 35–1 (2015).

40. Roberts, N., Grover, V.: Leveraging Information Technology Infrastructure to Facilitate a Firm's Customer Agility and Competitive Activity: An Empirical Investigation. J. Manag. Inf. Syst. 28, 231–270 (2012).
41. Sambamurthy, V., Bharadwaj, A., Grover, V.: Shaping Agility through Digital Options: Reconceptualizing the Role of Information Technology in Contemporary Firms. MIS Q. 27, 237–263 (2003).
42. Weiblen, T., Chesbrough, H.W.: Engaging with startups to enhance corporate innovation. Calif. Manage. Rev. 57, 66–90 (2015).
43. Carlile, P.R.: Transferring, Translating, and Transforming: An Integrative Framework for Managing. Organ. Sci. 15, 555–568 (2004).
44. Tushman, M.L., Katz, R.: External Communication and Project Performance. Manage. Sci. 26, 1071–1085 (1980).
45. Hansen, M.T., Løvås, B.: How do multinational companies leverage technological competencies? Moving from single to interdependent explanations. Strateg. Manag. J. 25, 801–822 (2004).
46. Cardinal, L.B., Turner, S.F., Fern, M.J., Burton, R.M.: Organizing for Product Development Across Technological Environments. Organ. Sci. 22, 1000–1025 (2011).
47. von Hippel, E., Katz, R.: Shifting Innovation to Users via Toolkits. Manage. Sci. 48, 821–833 (2002).
48. Fleming, L., Sorenson, O.: Technology as a complex adaptive system: Evidence from patent data. Res. Policy. 30, 1019–1039 (2001).

Why should Incumbent Firms jump on the Start-up Bandwagon in the Digital Era? – A Qualitative Study

Nihal Islam[1], Peter Buxmann[1], and Nicole Eling[1]

[1]TU Darmstadt, Information Systems, Darmstadt, Germany
{islam,buxmann,eling}@is.tu-darmstadt.de

Abstract. Due to ongoing digitalization, the traditional business models of incumbent firms are threatened by the innovation performance of start-ups. Therefore, a few incumbent firms have established programs to collaborate with start-ups in order to receive relevant impulses from them. However, empirically, there has barely been any insight into the specific role that start-ups play for incumbent firms. For this purpose, we present the key findings of our qualitative study, which has been built on interviews (n=35) with experts. Our results reveal that incumbent firms and start-ups have equally environmental factors affecting their decision-making in pursuit of digitalization. However, they differ in their technological and organizational factors. According to our findings, we have emphasized that collaboration between incumbent firms and start-ups could be an opportunity to meet the challenges of the digital era. They can build on these identified enabling factors of the partner and overcome their own inhibiting factors.

Keywords: Digitalization, Enabling and Inhibiting Factors, TOE Framework, Collaboration between Incumbent Firms and Start-ups, Qualitative Study

1 Introduction

Over recent years, new digital technologies have enabled lots of physical products and services to be turned into intangible digital content, such as the integrated usage of maintenance recommendations based on the vehicle data of a connected car in real time. Another example is the payment via a mobile wallet that promises infinitely more comfort in the daily life of consumers [1], [2]. The work environment has changed significantly as well. In particular, work models have become flexible and mobile, opened up by smartphones, tablets, and laptops so that individuals are no longer bound to any specific work space [3]. Last but not least, our present modes of communication have created a space for sharing and exchanging information on social media platforms [4].

The growing opportunities due to digital technologies – characterized by trends in information technology such as Social, Mobile, Analytics, and Cloud [5] – are forcing incumbent firms to rethink and realign their business models, and especially to change their operational processes and functional structures [6], [7]. Various conferences and expos focus on this issue, for example, Thinking Digital 2016 in the UK, Cebit 2016

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Islam, N.; Buxmann, P.; Eling, N. (2017): Why should Incumbent Firms jump on the Start-up Bandwagon in the Digital Era? - A Qualitative Study, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1378-1392

in Germany, the CEO2CEO Summit 2015 in the US, and even the annual meeting of the world economic forum in 2016. What is new in these discussions is that digital technologies have become the primary driver of innovation and reach the sphere of all industries and dissolve the market boundaries between industries. In particular, innovative products and services – especially those in digital form – have increasingly been placed in the center of consideration and are becoming a critical factor for success in the digital era [7], [8]. In addition, new entrants – especially start-ups from various industries – are represented in greater numbers than ever before. For instance, in 2015, the number of start-ups rose sharply, and it has been rising especially in Silicon Valley, London, and Berlin for the last several years [9]. Thereby, while start-ups are known as a major source of innovation because they use new technologies to invent products and services, and especially to develop digital innovations, incumbent firms are beginning to address the opportunities and challenges of digitalization [6], [10], [11], [12].

In the literature, the topic of digitalization in organizations is reflected in different ways. However, a holistic view of the connection between incumbent firms and start-ups in the context of digitalization has barely been investigated. As digital innovation is seen as crucial to the success of firms in creating new value for their business [8], a study is needed in order to analyze the pursuit of digitalization in incumbent firms and start-ups. Against this background, we have approached this field of interest with an observation about the causes of their opportunities and challenges in context of digitalization. In our study, we have used the adoption of digital technologies to create digital innovation as an outcome measure of digitalization. Information Systems (IS) research has broadly researched technology adoption at the individual and organizational levels. However, none of these studies have explained the adoption of digital technologies with specific relevance to creating digital innovation by comparing incumbent firms with start-ups. By taking these observations together, the addressed research question is: ***What factors influence the decision to adopt digital technologies in order to create digital innovations by incumbent firms and start-ups?*** To generate insights into this emergent phenomenon of digitalization and an understanding of the different perspectives of experts, we conducted an exploratory study based on interviews with 23 executive managers (most of them CEOs and CIOs) from various incumbent firms, as well as with 12 founders of start-ups. The objective of this study was to retain the richness of the phenomenon while studying the linkage between known but less researched factors of incumbent firms´ and start-ups´ decision-making in the context of digitalization. Thereby, the variety of the identified factors was sorted by the technology-organization-environment (TOE) framework, which is suited to our understanding of the affects of these factors. In particular, based on our findings, we have illustrated what role start-ups play for incumbent firms and why collaboration between both is a great opportunity to meet the challenges of the digital era.

The paper is structured as follows: First, we provided a brief overview of the theoretical background and related work to mark off the research field. Then, we described how the exploratory study was designed and how the interviews were executed. Thereby, we have presented our sample of 23 executive managers of incumbent firms and 12 founders of start-ups. In addition, we have presented the empirical results and integrated our findings by utilizing the TOE framework in a sorted form. Therefore, the results of our paper provide a holistic view of the enablers and inhibitors that are relevant for incumbent firms and start-ups in the digital era. Afterwards, we discussed our key findings and illustrated an approach to future work. Finally, we concluded by describing the limitations of and contributions to research and practice.

2 Theoretical Background and Related Work

The sociotechnical process of applying new technologies to broader social and institutional contexts comprises the term "digitalization" [2], [13]. Currently, there is a new wave of digitalization because the growing role of digital technologies is changing the way firms relate to their customers [14]. The objective of today's discourse on digitalization is not only to improve efficiency based on new technologies, but also to create new business value with innovative products, services, or business models – especially in the form of digital innovation, which is embodied in or enabled by IT [8], [12]. In particular, SMAC can serve as a holistic basis and equip an organization to create digital innovations in context of digitalization [5]. Thereby, an organizational innovation is defined as the first use of an idea in terms of a product, process, or service that is new or improved to the organization adopting it [15], [16], [17].

In IS research, there have been some studies conducted and empirical evidence found for the impacts of digitalization on organizations. For instance, IS researchers have investigated the IT-enabled transformational change in organizations, the importance of a digital business strategy, changes in the producer-consumer relationship, and the managerial tasks of a chief digital officer [6], [7], [18], [19]. In addition, there are concepts for classifying organizations into types of digital maturity levels [e.g. 20]. However, research in the field of digitalization with a focus on whether incumbent firms and start-ups – despite culture clashes – fit together well can hardly be found. Only a few studies exist that have made collaboration between incumbent firms and start-ups a subject of discussion; however, the specific role of start-ups for incumbent firms has not been investigated in this context. Rather, it has been found by researchers that, when both work together, it is a balancing act, as it is a cooperation and competition at the same time. In addition, it is also about strategies for dealing with emerging cooperative competitive tensions [21]. Although some incumbent firms have a great interest in cooperating and although they have great advantages in open innovation [22], [23]; another study has resulted in a more differentiated result: The choice of the wrong cooperation partner does not promote the innovation of both sides. Above all, startups, by and large, are dissuaded from cooperation with multinational firms [24].

To illustrate the specific characteristics of start-ups related to incumbent firms, the life-cycle approach could be used. This approach is based on the assumption of the ideal-type stages of a firm: introduction, growth, maturity, and decline. Start-ups can be categorized into the introduction and growth stages. These stages characterize high investment costs via extraction of necessary resources, which usually causes low turnover [25]. Besides that, there are lots of start-ups that have very low investment costs, as they only sell digitized products or services without huge production costs.

Generally, start-ups are not bound to a specific industry. However, a distinguishing characteristic of start-ups can be the degree of innovation. An innovative start-up has a solid foundation for creating technology-specific innovation. In particular, high flexibility and digital know-how enable start-ups to quickly implement ideas as the innovative digital products and services [26], [27], [28]. In the growth stage, firms successfully penetrate the market. The growing size of firms indicates the standardization and professionalization of all the operational systems and processes. Incumbent firms are located at the end of this growth stage when reaching the maturity stage [25]. Thereby, incumbent firms are characterized by their good position in the market [29]. Despite incumbent firms and start-ups differing on characteristics, it can be assumed that both address the opportunities and challenges of digitalization [30]. With this in mind, a holistic view of the factors influencing the adoption of digital technologies in consideration could be useful for understanding the decision-making in the context of digitalization from the two parties' perspectives. A broadly used framework in the field of organizational technology adoption is the TOE framework. Accordingly, technological innovation decision-making is influenced by technological development, organizational, and environmental dimensions. The TOE provides a set of factors that are relevant: The technological context includes the availability and characteristics of technologies. Factors in the organizational context are formal and informal linking structures, communication processes, firm size, and slack. The environmental context describes factors outside of the organization. These include in particular industry characteristics and market structure, technology support infrastructure, and government regulation [31]. Previous studies provide relevant factors as well [32], [33], [34], [35], [36], [37]. However, a holistic view of the factors that influence a specific kind of organization's adoption, such as that which compares incumbent firms with start-ups directly of a particular technological innovation (e.g., digital innovation), is missing. Furthermore, to our knowledge, existing studies have not focused on incumbent firms and start-ups with a holistic view or with the aim to analyze how these different factors of incumbent firms and start-ups could fit together. Following this line of thought, the applicability of the TOE framework supports our research by integrating our findings in a sorted form.

3 Research Study

The qualitative study presented here was a research project that investigated effects on and outcomes of digitalization in business and science. In this paper, we have presented relevant factors influencing decision-making in the context of digitalization – identified by statements from incumbent firms and start-ups – with the aim to gain insights about the role of start-ups for incumbent firms.

As described earlier, this study is focused on a relatively new phenomenon. Against this background, we decided on qualitative research with an explorative design. The explorative approach allowed us to analyze data material in areas in which only limited knowledge exists [38], [39]. Our qualitative study has been built on interviews with experts. Generally, an expert is a person with special knowledge of a subject area [40]. For the interviews, a semi-structured guideline was used with questions that were designed to generate comparability of results and were selected in order to preserve the exploratory character [38]. This guideline ensured that all interviews covered the main topic, and it allowed us to address the peculiarities of the respective firms' contexts.

In the following, the sample and data collection process, as well as the data analysis method, are described in detail.

3.1 Sample and Data Collection Process

The focus of this paper is to present the differences between incumbent firms and start-ups by identifying the influencing factors for adopting digital technologies in the context of digitalization.

Interviewees were first asked about the importance of digitalization in their organizations, then about the activities in the areas of projects in terms of digital innovations, strategy, processes, leadership style, and culture. Thereby, the interviewees would give insights into their own field of activities, as well as a holistic view of their organization across all departments.

Firstly, we concentrated on interviews with executive managers (most of them CEOs and CIOs) of well-known and well-established firms from various industries, such as consulting, product-oriented, and service-oriented firms. With the intention of developing a uniform and industry-nonspecific picture of firms, experts were selected from various industries. Within the incumbent firms, we decided to interview managers from the strategic level because they determine the strategy of the firm and have a holistic cross-functional organizational overview. This view is crucial, as digitalization affects all functions of a firm [6], [7]. In addition to these interviews with managers of incumbent firms, we conducted interviews with founders of various start-ups. The start-ups interviewed were required to have technology-based business models with the focus on business-to-business (B2B) and business-to-consumer (B2C). This was because of the assumption that they would demonstrate a high degree of innovation performance in context of B2B and B2C [10], [11].

The expert interviews were conducted during the period between June 2015 and August 2015. In total, the sample was comprised of 35 interviews with 23 executive managers (hereby abbreviated as "IF" for "incumbent firm") and 12 founders of start-ups (hereby abbreviated as "SU" for "start-up"). All the interviewees from the incumbent firms had a proactive role and extensive staffing/budget responsibility within their firms at the time of the interview. Beyond that, all the founders of start-ups interviewed had been managing their business for at least one year. Table 1 provides an overview of the 35 experts interviewed. Thereby, we ranked each group by founding year, because, in all likelihood, the older the organization, the more established the existing business model.

Table 1. Overview of experts interviewed

Group IF: Executive managers of incumbent firms									
ID	**Respondent's Position**	**Firm Sector**	**Founding Year**	**Employees (in 2015)**	**ID**	**Respondent's Position**	**Firm Sector**	**Founding Year**	**Employees (in 2015)**
IF-01	CIO	Banking & Finance	1870	101.104	**IF-13**	CEO	Consulting	1992	280
IF-02	CEO	Manufacturing	1895	5.600	**IF-14**	Senior Vice President	Transport	1994	300.000
IF-03	Board IT	Insurance	1922	14.505	**IF-15**	Head of Global IT Enterprise	Pharmaceutical	1995	39.639
IF-04	CEO	Manufacturing	1945	5.700	**IF-16**	CEO	Enterprise Software	1998	400
IF-05	Division Manager	Transport	1947	20.720	**IF-17**	CEO	Software	1998	250
IF-06	CIO	Transport	1953	119.559	**IF-18**	Division Manager IT	Energy Supplier	1998	2.732
IF-07	Regional CEO	Manufacturing	1953	6.831	**IF-19**	Senior Vice President IT	Consulting	2000	45.990
IF-08	CEO	Enterprise Software	1969	4.421	**IF-20**	Executive Partner	Consulting	2000	60
IF-09	Senior Vice President	Enterprise Software	1972	76.986	**IF-21**	CEO	Enterprise Software	2000	50
IF-10	CEO	Manufacturing	1976	1.120	**IF-22**	CEO	Consulting	2002	65
IF-11	Chairman of the Board	Consulting	1989	400.000	**IF-23**	CEO	Consulting	2007	50
IF-12	Managing Partner	Consulting	1989	212.000					
Group SU: Founders of start-ups									
ID	**Business Field**	**Business Model Orientation**	**Founding Year**	**Employees (in 2015)**	**ID**	**Business Field**	**Business Model Orientation**	**Founding Year**	**Employees (in 2015)**
SU-01	IT Security	B2C	2012	3	**SU-07**	Big Data	B2B	2014	6
SU-02	Legal Tech	B2B	2013	4	**SU-08**	Mobile App	B2B	2014	4
SU-03	Online Marketing	B2B	2013	14	**SU-09**	IT Security	B2B	2015	4
SU-04	Digital Printing	B2C	2014	2	**SU-10**	IT Security	B2B	2015	2
SU-05	Online Recruiting	B2B	2014	24	**SU-11**	IT Energy	B2C	2015	2
SU-06	Online Training	B2B	2014	10	**SU-12**	Mobile App	B2C	2015	6

The interviews were held in private spaces and lasted an average of 45 minutes. All interviews were recorded. For easier analysis, the recorded material was transcribed. This process resulted in 309 DIN A4 format pages of transcripts.

3.2 Data Analysis Method

The aim of the data analysis was to retain and provide essential contents by abstracting a manageable collection of data that still illustrated a reflection of the data material. Characteristic of this type of examination is the methodological technique "content analysis" [41]. Thereby, we have used an inductive approach, as we have not had theoretical assumptions in context of our research. Against this background, the categories were derived inductively from the transcribed interviews and, thus, were not predefined or derived from existing theory. Based on the content analysis technique and following the reducing code rules, the data material was reduced into an abstract form in order to paraphrase and generalize the data material by maintaining only the parts of substantial content, which was finally divided into categories [42], [43]. For instance, the quotation of an expert *"Each new project, such as one based on digitalization to create digital innovation, always presupposes a well-realized application platform. […] we have a well-functional basis that provides a functional IT System within our firm."* (IF-19)" was coded – after a paraphrasing and generalizing process – to category 'Solid IT Infrastructure'. As required, corresponding points in the material were assigned to the newly formed categories.

To achieve reliability in our analysis, multiple people (three in total) coded and analyzed the data material by using a software tool [44]. Thereby, we have combined all the categories together and marked only those that were coded by all. Afterwards, we have filtered them by factors that are relevant in each respective incumbent firm and start-up. Finally, the categories were filtered again by relevance in terms of representing insights into factors influencing incumbent firms´ and start-ups´ decision-making in context of digitalization.

4 Empirical Results

To illustrate our key findings in a comprehensive view, the relevant categories were summarized into thematically related groups. For instance, the categories ‘solid IT infrastructure’ and ‘access to broadband’ were summarized into **‘Available Technical Equipment’**. As a result, these groups stand for 15 influencing factors for each respective incumbent firm and start-up that exist across all industries. In particular, we found that factors can be enablers or inhibitors and differ partially, depending on the incumbent firm or start-up. Thereby, enablers help to promote the adoption of digital technologies, while inhibitors prevent their adoption [45]. It is possible that the same factors may be rated differently depending on the organization in which they occur, as is presented in our study.

Due to the variety of the identified factors, in the following, the influencing factors have been sorted into an aggregated form based on the TOE framework. Accordingly, the factors were classified as technology, organization, and environment, whereby these mutually influence each other. As illustrated in Figure 1, the factors of incumbent firms and start-ups have been sorted by enabling factors (the symbol “+”) and inhibiting factors (the symbol “-”). Thereby, the symbol “+” means that the identified factor has an enabling effect, and the symbol “-” means that the factor has an inhibiting effect on the decision-making of incumbent firms and start-ups in the context of digitalization.

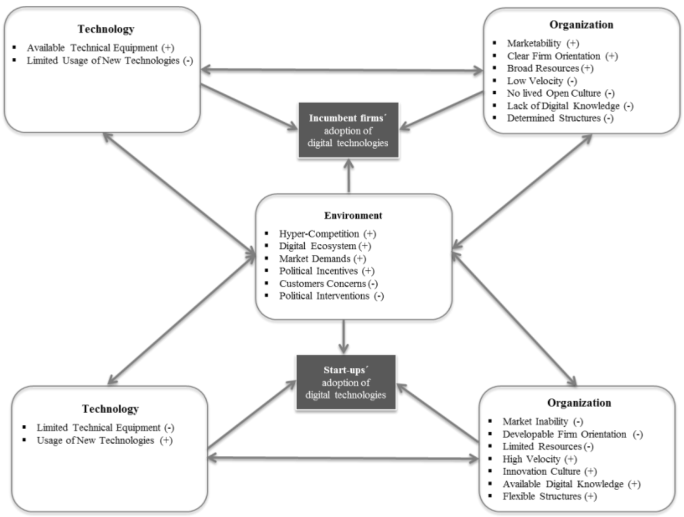


Figure 1. Influencing factors for decision-making in the context of digitalization

The factors illustrated above mark the enablers and inhibitors that pave the way for understanding the similarities and differences between incumbent firms and start-ups. As the factors originate from the statements from the interviews, the definitions of the factors can be derived from the explanations of the experts interviewed. In the following, we have described some examples of each technological, organizational, and environmental factor in incumbent firms and start-ups.

The technological context includes the internal and external technologies that are available for an organization and that fit with an organization´s current technology. For many IT solutions to work, firms need **'Available Technical Equipment'**, which allows the performing of digital-based projects. Some experts from incumbent firms stated that a 'solid IT infrastructure´ is a highly relevant enabler for them to operate without problems. Furthermore, 'access to a broadband connection' is a basic prerequisite for and enabler of efficient use of modern IT solutions in the digital era. In particular, for projects to develop digital products and services, available technical equipment is needed in order to build upon on this foundation. In contrast, **'Limited Technical Equipment'** causes start-ups to be restricted in their business actions. Some start-up interviewees stated that their 'IT infrastructure was not comprehensive` and hindered them in implementing further digital solutions based on existing ideas. Furthermore, some founders of start-ups stated that a 'reduced broadband connectivity' led to fewer additional digital services within their business models.

The organizational context encompasses the characteristics and resources of an organization. Thereby, **'Marketability'** is an essential factor for incumbent firms. This means that 'capability due to experience in business' enables incumbent firms to operate in the market with perseverance. Moreover, they have an understanding of environmental dynamics. Furthermore, 'product and service portfolio' is an enabler that represents a solid foundation for extending functional business to a digital business model. However, start-ups are not able to act in their business field due to '**Market Inability'**. This is attributable to their 'lack of experience in business' and 'difficult access to customers', as mentioned by some founders of start-ups interviewed.

The environmental context includes, among others, the competition pressure in the course of ongoing digitalization. Thereby, **'Hyper-Competition'** especially leads to the rapid escalation of competitive tactics. Against this background, the executive managers of incumbent firms indicated that there were 'new entrants from various industries'. The founders of start-ups used the same line in relation to 'other start-ups'. Both indicated that hyper-competition enables the adoption of digital technologies, as it is crucial to keep a competitive advantage.

For a closer look at the results, we compared the factors of incumbent firms and start-ups and analyzed our key findings with a focus on the role of start-ups for incumbent firms. In the following, the factors are represented with sub-factors in order to highlight the relevance – stated by the interviewees – of the enabling and inhibiting factors. Thereby, relevance has been considered based on the following point-allocation: ■ = very high relevance (p=4), ▣ = high relevance (p=3), ⊡ = low relevance (p=2), and ☐ = very low relevance (p=1). The findings have been rated according to the following formula: $\sum$**p/n**, with n=number of mentions per executive manager with respect to start-ups. In Table 2, the average relevance is shown for each respective factor.

Table 2. Merging both groups*

Incumbent Firms		Incumbent factor	Incumbent category		Start-up category	Start-up factor		Start-ups
Technology								
Enabler	■	Solid IT Infrastructure	**Available Technical Equipment**	↔	**Limited Technical Equipment**	No comprehensive IT Infrastructure	■	Inhibitor
Enabler	▣	Access to Broadband Connection				Reduced Broadband Connection	■	Inhibitor
Inhibitor	■	Legacy IT Systems	**Limited Usage of New Technologies**	↔	**Usage of New Technologies**	IT-Based Innovative Tools	■	Enabler
Inhibitor	▣	No Broad Utilization of Cloud Computing				Full Utilization of Cloud Computing	■	Enabler
Organization								
Enabler	■	Capability Due to Experience in Business	**Marketability**		**Market Inability**	Lack of Experience in Business	▣	Inhibitor
Enabler	■	Product and Service Portfolio				Difficult Access to Customer	■	Inhibitor
Enabler	■	Digital Business Strategy	**Clear Firm Orientation**	↔	**Developable Firm Orientation**	Strategic Goal Still in Development	■	Inhibitor
Enabler	☐	Process Orientation				Non-professional Flowcharts	☐	Inhibitor
Enabler	▣	Comprehensive Workforce	**Broad Resources**		**Limited Resources**	Mini Manpower	⊡	Inhibitor
Enabler	⊡	Available Budget				Scarce Budget	■	Inhibitor
Inhibitor	■	Hierarchical Structures	**Low Velocity**		**High Velocity**	Flat Organization	■	Enabler
Inhibitor	▣	High Quality Requirements				No Long Discussions	▣	Enabler
Inhibitor	■	Longstanding Methods				Agile Working Methods	■	Enabler
Inhibitor	▣	Missing Entrepreneurial Spirit	**No lived Open Culture**		**Innovation Culture**	Willingness to Dare	■	Enabler
Inhibitor	⊡	Older Top Management		↔		Trial-and-Error Attitude	■	Enabler
Inhibitor	■	Aging Workforce	**Lack of Digital Knowledge**		**Available Digital Knowledge**	Digital Natives	■	Enabler
Inhibitor	▣	Limited Experiences in Big Data				Deal with Big Data	▣	Enabler
Inhibitor	■	Low Availability of IT Professionals						Enabler
Inhibitor	■	Inefficent Interaction of Business and IT	**Determined Structures**		**Flexible Structures**	Jack of All Trades	■	Enabler
Inhibitor	▣	Fixed Workplaces				Co-working Spaces	▣	Enabler
Environment								
Enabler	■	New Entrants from Various Industries	**Hyper-Competition**			Other Start-ups	■	Enabler
Enabler	■	Knowledge Exchange with Start-ups	**Digital Ecosystem**			Knowledge Exchange with Other Start-ups	■	Enabler
Enabler	☐	Platforms for Digitalization Topics				Platforms for Start-ups	▣	Enabler
Enabler						Venture Capital Investors	■	Enabler
Enabler	■	Customer Expectations	**Market Demands**			User Entrepreneur	■	Enabler
Enabler	☐	Funding Programs	**Political Incentives**			Funding Programs	⊡	Enabler
Inhibitor	■	IT Security Concerns	**Customers Concerns**			IT Security Concerns	■	Inhibitor
Inhibitor						No Openess Due to Unknown Name	▣	Inhibitor
Inhibitor	■	Law Protection of Data	**Political Interventions**			Law Protection of Data	■	Inhibitor
Inhibitor	⊡	Regulatory Requirements				Bureaucratic Expenditure	⊡	Inhibitor

**A list of all the factors, with corresponding descriptions, is available from the authors upon request.*

As seen in Table 2, incumbent firms and start-ups share equal environmental factors in the digital era. For instance, interviewees from both positions stated that **'Hyper-Competition'** is a great enabler for promoting decision-making in the context of digitalization in their own firms. Aside from that, they had been facing 'Customer Concerns', which inhibit incumbent firms and start-ups equally. However, although the environmental factors of incumbent firms and start-ups are equivalent, the factors of technology and organization differ significantly. Thereby, the technological and organizational enabling factors identified in incumbent firms are similar to the inhibiting factors of start-ups, and, likewise, the inhibiting factors of incumbent firms are similar to the enabling factors of start-ups. For instance, the organizational enabling factor **'Marketability'** for incumbent firms is missing in start-ups and can be compared with the inhibiting factor **'Market Inability'**. Beyond this, there is also the indication that the inhibitors of incumbent firms can correspond directly with the enablers of start-ups. For instance, the technological inhibiting factor **'Limited Usage of New Technologies'** for incumbent firms can be compared with **'Usage of New Technologies'** for start-ups: While incumbent firms would have liked to implement a higher level of cloud computing if they had not had concerns about their data, start-ups did not have concerns about using cloud computing solutions.

5 Discussion of Findings

To illustrate and discuss the new findings of our research, we have compared our results with those of existing studies. We have drawn on relevant studies with respect to the specific context of the organizational adoption of technological innovation, which can be associated with digital technologies. These include in particular adoption of cloud computing, customer-based interorganizational system (IOS), e-business, e-commerce, and mobile business. An overview of the relevant factors is provided in Table 3.

Table 3. Overview of the relevant factors

		Innovation
Technological factors	Complexity [32], [35]	Cloud Computing
		Customer-Based IOS
	Compatibility[32], [35]	Cloud Computing
		Customer-Based IOS
	Technology Competence (e.g., IT-Infrastructure) [33], [36], [34]	E-Business
		E-Commerce
	Unresolved Technical Issue [36]	E-Commerce
	Interoperability [36]	
Organizational factors	Top-Management Support [32]	Cloud Computing
	IS Experience [32]	
	Firm Scope [33], [34]	E-Business
	Top Management Support [36], [35]	Customer-Based IOS
		E-Commerce
	Strategy in terms of Technology [36]	E-Commerce
	Cost-Benefit Assessment [36]	
	Financial Commitment [34]	E-Business
	Managerial Obstacles [37]	Mobile Business
Environmental factors	Competitive Pressure [32], [33], [34], [35], [37]	Cloud Computing
		E-Business
		Customer-Based IOS
		Mobile Business
	Trading Partner Readiness [33]	E-Business
	Consumer Readiness [33]	
	Customer Interaction [35]	Customer-Based IOS
	Legal Issue [36]	E-Commerce
	Regulatory Support [34]	E-Business

In a comparison of the factors presented above with our findings, it can be noted that there is at least one newly identified factor in each dimension of the TOE separated between incumbent firms and start-ups: In the technology context, the factors **'Limited Usage of New Technologies'** in incumbent firms and **'Usage of New Technologies'** in start-ups have been identified; **'Marketability'**, **'Low Velocity'**, **'No lived Culture',** and **'Determined Structures'** in incumbent firms and **'Market**

Inability', **'High Velocity'**, **'Innovation Culture'**, and **'Flexible Structures'** in start-ups have been identified as organizational factors; and **'Digital Ecosystem'** in both represents an environmental factor. The rest of our findings can be found in a similar form in existing literature.

Besides the new identified factors, based on our key findings, we have been able to investigate whether incumbent firms should collaborate with start-ups in the digital era. Basically, the bigger a firm, the higher the probability – despite sufficient relevant resources, such as budget and technical equipment – that they will have lower innovative performance, due to, for example, inertia [29], [46]. In contrast, start-ups have relevant enablers for adopting digital technologies to create digital innovation, which are important in the digital era. However, incumbent firms have several inhibiting factors for adopting digital technologies to create digital innovation; hence, collaboration with start-ups can be an opportunity to balance these factors. In addition, incumbent firms also have relevant enabling factors that start-ups need for expanding their efforts in the field of digital innovations.

It has become visible that incumbent firms and start-ups complement each other perfectly and have the potential to collaborate with each other in an appropriate form. Against this background, the approach of open innovation can be an option for initiating collaborations between incumbent firms and start-ups. This potential was also emphasized by some of the executive managers of the incumbent firms: "*We are definitely interested in working with start-ups"* (IF-18). Start-ups are seen as providing a chance to create innovative products and services, as pointed out by one interviewee in following words: *"It is a great opportunity for large firms to work with start-ups – you can find your missing impulses right there"* (IF-13). These and other quotations paint a picture of incumbent firms being interested in collaborating with start-ups, as well as hoping to learn from them. In particular, incumbent firms can use start-ups to create an entrepreneurial organization with the aim to stimulate the expansion of competence and, across all departments, to build up the ability to act and operate entrepreneurially. Moreover, with collaboration with start-ups, new business areas can be pursued [47], [48].

However, when it comes to a concrete activity on which to collaborate with start-ups, start-ups have criticized incumbent firms' restricted openness towards collaboration. One founder of a start-up described it this way: *"Cooperation partners are very important for achieving drive in our target market. However, discussions with potential partners do not lead to positive results"* (SU-11). In the same line of thinking, some founders of start-ups stated that incumbent firms had *"difficulties with cooperation because they do not understand our innovative product and services"* (SU-02). This reproach could be explained by the assumption that *"potential cooperation partners do not really see the potential of our business model"* (SU-04). To sum it up, while some incumbent firms want to collaborate with start-ups, there are apparently some noticeable hurdles. This conflicts with some popular views and raises a fundamental question concerning the usefulness of collaboration between incumbent firms and start-ups to pursuing increased innovation performance on both sides. Currently, the design of corporate start-up programs is being intensely examined, especially with regard to the joint development of new and innovative ideas [23]. At the same time, many programs are promising success, in which the start-ups should profit from the experience and resources of the incumbent firms.

Often, however, studies on this topic have mainly been based on the experience of the incumbent firms, or on the prerequisites and recommendations for successful collaboration for managers of incumbent firms [22].

6 Conclusion, Limitation, and Outlook

On the grounds of reflection on the enabling and inhibiting factors influencing the incumbent firms´ and start-ups´ decision-making in the context of digitalization, we have been able – in our opinion – to present newly identified influencing factors that were sorted by the TOE framework. Alongside this, we have analyzed the similarities and differences between incumbent firms and start-ups: In particular, we have found out that the technological and organizational enabling factors in incumbent firms are similar to the inhibiting factors in start-ups, and the corresponding inhibiting factors in incumbent firms are similar to the enabling factors in start-ups. In consideration of how inhibitors can develop into enablers, incumbent firms as well as start-ups can approach each other to overcome their respective inhibiting factors. There are already a few incumbent firms trying to establish structures in which to collaborate with start-ups. For instance, corporate-startup programs, namely, accelerator programs like the accelerator program "Microsoft Ventures" and the newly founded "InnoJam++" event from SAP in cooperation with Volkswagen. However, there are lots of incumbent firms that have not yet recognized the opportunity of cooperating with start-ups to meet their challenges in the digital era. Against this background, we assume that our results could have an impact on the reinforcement of incumbent firms' willingness to collaborate with start-ups.

As in any study, our qualitative research has been constrained by some limitations. However, at the same time, these limitations provide avenues for further research. Due to the interpretive nature of our research, the results we have described represent the sense-making process of the researchers. Subjective personal judgments cannot be ruled out completely, even though we took great care to reflect the subjects' opinions as correctly as possible. Moreover, the factors have been derived from the views stated by the interviewees. It cannot be ruled out that there are more factors that we have not identified in our study. Besides that, it is difficult to make quantitative predictions. Therefore, it is necessary to validate our results with an extensive investigation based on a quantitative study. Nevertheless, despite the limitations, our study makes three major contributions: *Firstly*, we have illustrated influencing factors that are relevant to incumbent firms´ and start-ups´ decision-making in pursuit of digitalization. In particular, we have provided new findings related to enabling and inhibiting factors from a holistic viewpoint as a basis for the research discussion. *Secondly*, for practitioners, we have shown the potential of start-ups as cooperation partners and have emphasized that incumbent firms should collaborate with start-ups in order to be competitive in the digital era. *Thirdly*, we have indicated that hurdles exist in efforts to collaborate, which should be investigated in detail in future research. Against this background, it would be interesting to analyze the following main question: What factors, and how do those factors, influence collaboration between incumbent firms and start-ups in the context of the digital era? Thereby, the dyadic relationship should be considered in this investigation, with the primary aim of figuring out what factors are relevant from a two-sided perspective, as our results have illustrated that there remains an area of tension. Thus, it is advisable to examine precisely which hurdles exist in order to determine a balance that can be crucial for the success of the collaboration between incumbent firms and start-ups.

References

1. Yoo, Y.: Digitalization and Innovation. Institute of Innovation Research, Hitotsubashi University (2010)
2. Yoo, Y.: Computing in everyday life: A call for research on experiential computing. Mis Quarterly 34, 213-231 (2010)
3. Serrano, C., Boudreau, M.-C.: When Technology Changes the Physical Workplace: The Creation of a New Workplace Identity. (2014)
4. Obar, J.A., Wildman, S.S.: Social Media Definition and the Governance Challenge: An Introduction to the Special Issue. (2015)
5. Caldwell, F.: SMAC in the Middle of the Nexus at LegalTech. Gartner Blog Network (2013)
6. Matt, C., Hess, T., Benlian, A.: Digital transformation strategies. Business & Information Systems Engineering 57, 339-343 (2015)
7. Bharadwaj, A., El Sawy, O.A., Pavlou, P.A., Venkatraman, N.V.: Digital business strategy: Toward a next generation of insights. Mis Quarterly 37, 471-482 (2013)
8. Fichman, R.G., Dos Santos, B.L., Zheng, Z.E.: Digital Innovation as a Fundamental and Powerful Concept in the Information Systems Curriculum. Management Information Systems Quarterly 38, 329-353 (2014)
9. Compass: The Global Startup Ecosystem Ranking 2015. (2016)
10. Verhees, F.J., Meulenberg, M.T.: Market orientation, innovativeness, product innovation, and performance in small firms. journal of small business management 42, 134-154 (2004)
11. Ackermann, S.J.: Are small firms important? Their role and impact. Springer Science & Business Media (2012)
12. Fitzgerald, M., Kruschwitz, N., Bonnet, D., Welch, M.: Embracing Digital Technology. MIT (2013) (2013)
13. Tilson, L.a.S.: Research Commentary—Digital Infrastructures: The Missing IS Research Agenda. Information Systems Research 21, 748-759 (2010)
14. Piccinini, E., Gregory, R.W., Kolbe, L.M.: Changes in the Producer-Consumer Relationship-Towards Digital Transformation. Wirtschaftsinformatik, pp. 1634-1648 (2015)
15. Daft, R.L.: A dual-core model of organizational innovation. Academy of management journal 21, 193-210 (1978)
16. Barnett, H.G.: Innovation: the basis of cultural change. McGraw-Hill, New York, NY, US (1953)
17. Rogers, E.M.: The innovation-decision process. Diffusion of innovations 5, 168-218 (1983)
18. Horlacher, A., Hess, T.: What Does a Chief Digital Officer Do? Managerial Tasks and Roles of a New C-Level Position in the Context of Digital Transformation.

2016 49th Hawaii International Conference on System Sciences (HICSS), pp. 5126-5135. IEEE (2016)
19. Lucas Jr, H.C., Agarwal, R., Clemons, E.K., El Sawy, O.A., Weber, B.W.: Impactful Research on Transformational Information Technology: An Opportunity to Inform New Audiences. Mis Quarterly 37, 371-382 (2013)
20. Westerman, G., Tannou, M., Bonnet, D., Ferraris, P., McAfee, A.: The Digital Advantage: How digital leaders outperform their peers in every industry. MITSloan Management and Capgemini Consulting, MA 2-23 (2012)
21. Ansari, S.S., Garud, R., Kumaraswamy, A.: The disruptor's dilemma: TiVo and the US television ecosystem. Strategic Management Journal (2015)
22. Kohler, T.: Corporate accelerators: Building bridges between corporations and startups. Business Horizons 59, 347-357 (2016)
23. Lisowska, R., Stanisławski, R.: The Cooperation of Small and Medium-sized Enterprises with Business Institutions in the Context of Open Innovation. Procedia Economics and Finance 23, 1273-1278 (2015)
24. Antolin-Lopez, R., Martinez-del-Rio, J., Cespedes-Lorente, J.J., Perez-Valls, M.: The choice of suitable cooperation partners for product innovation: Differences between new ventures and established companies. European Management Journal 33, 472-484 (2015)
25. Heinrichs, N.: Bewertung von Wachstums-und Startup-Unternehmen. Diplomica Verlag (2008)
26. Reis, E.: The Lean Startup: How constant innovation creates radically successful businesses. Portfolio Penguin, US (2011)
27. Colombelli, A., Krafft, J., Vivarelli, M.: To be born is not enough: the key role of innovative start-ups. Small Business Economics 1-15 (2016)
28. Hanks, S.W., C.; Jansen, E.; Chandler, G.: Tightening the Life-Cycle Construct: A Taxonomic Study of Growth Stage Configurations in High-Technology Organizations. Entrepreneurship, theory and practice : ET & P 18, 5-30 (1994)
29. Christensen, C.: The innovator's dilemma: when new technologies cause great firms to fail. Harvard Business Review Press (2013)
30. BarNir, A., Gallaugher, J.M., Auger, P.: Business process digitization, strategy, and the impact of firm age and size: the case of the magazine publishing industry. Journal of Business Venturing 18, 789-814 (2003)
31. Tornatzky, L.G., Fleischer, M., Chakrabarti, A.K.: Processes of technological innovation. Lexington Books (1990)
32. Alshamaila, Y., Papagiannidis, S., Li, F.: Cloud computing adoption by SMEs in the north east of England: A multi-perspective framework. Journal of Enterprise Information Management 26, 250-275 (2013)
33. Zhu, K., Kraemer, K.L., Xu, S.: Electronic Business Adoption by European Firms: A Cross-country Assessment of the Facilitators and Inhibitors. (2003)
34. Zhu, K., Kraemer, K.L.: Post-adoption variations in usage and value of e-business by organizations: cross-country evidence from the retail industry. Information systems research 16, 61-84 (2005)
35. Grover, V.: An empirically derived model for the adoption of customer- based interorganizational systems. Decision sciences 24, 603-640 (1993)

36. Teo, T.S., Ranganathan, C., Dhaliwal, J.: Key dimensions of inhibitors for the deployment of web-based business-to-business electronic commerce. IEEE Transactions on engineering Management 53, 395-411 (2006)
37. Picoto, W., Belanger, F., Palma-dos-Reis, A.: Leveraging on Mobile Business to Enhance Firm Performance: an Organizational Level Study. ECIS 2012, (2012)
38. Yin, R.K.: Case Study Research: Design and Methods. Sage, Los Angels, London, New Delhi, Singapore, Washington DC (2008)
39. Neuendorf, K.A.: The content analysis guidebook. Sage (2002)
40. Bogner, A., Littig, B., Menz, W.: Interviewing experts. Palgrave Macmillan Basingstoke (2009)
41. Weber, R.P.: Basic content analysis. Sage (1990)
42. Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. Beltz, Weinheim/Basel (2010)
43. Corbin, J., Strauss, A.: Basics of qualitative research: Techniques and procedures for developing grounded theory. Sage publications (2014)
44. Richards, L.: Handling qualitative data: A practical guide. Sage (2014)
45. Cenfetelli, R.T.: Inhibitors and enablers as dual factor concepts in technology usage. Journal of the Association for Information Systems 5, 16 (2004)
46. Ansari, S., Krop, P.: Incumbent performance in the face of a radical innovation: Towards a framework for incumbent challenger dynamics. Research Policy 41, 1357-1374 (2012)
47. Covin, J.G., Miles, M.P.: Corporate entrepreneurship and the pursuit of competitive advantage. Entrepreneurship: Theory and practice 23, 47-47 (1999)
48. Morris, M.H., Kuratko, D.F., Covin, J.G.: Corporate entrepreneurship & innovation. Cengage Learning (2010)

How to Design Patterns in IS Research – A State-of-the-Art Analysis

Andreas Günther[1] and Robin Knote[2]

[1] University of Kassel, Kassel, Germany
andreas.guenther@gumail.net
[2] University of Kassel, Information Systems, Kassel, Germany
robin.knote@uni-kassel.de

Abstract. Patterns are becoming increasingly prominent in the field of Information Systems (IS). They contain good practice solutions to recurring problems and are therefore especially valuable for systems development. Although a huge amount of patterns for nearly every problem in systems development (and even in many fields of IS such as enterprise information management or security) exist, literature still misses a systemized overview on how such patterns are developed. Approaches in practice exist but are either methodologically imprecise or lack a scientific foundation. This contribution is devoted to review the state-of-the-art on how patterns are designed. The findings of our systematic literature review reveal approaches for pattern development which we consolidate, structure and critically reflect. Since we argue that patterns can be considered design artifacts, we apply a design science research (DSR) lens to these approaches, highlight potential gaps and show needs for future development in theory and practice.

Keywords: pattern, pattern development, sociotechnical systems engineering, design science research

1 Introduction

The ongoing digital transformation causes great technological and societal challenges and opportunities. Consequently, the complexity of systems engineering is exponentially increasing. On the technical side, systems have to be able to communicate with an unknown number of different systems and subsystems, offer individualized services to the user and even change their behavior during runtime based on context information. On the societal side, systems have to be accepted by the users, which require privacy, trust, usability or legal considerations as well as the offering of added value. Therefore, the main challenges for systems engineering are identified as complexity, multidisciplinarity and user focus. Accordingly, systems development, in order to be efficient and effective, has to include heterogeneous and multidisciplinary stakeholders to analyze their requirements and offer proper designs [1]. A profitable approach for eliciting, understanding and applying multidisciplinary knowledge and experience within systems engineering are patterns [2–5]. Patterns

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Günther, A.; Knote, R. (2017): How to Design Patterns in IS Research – A State-of-the-Art Analysis,
in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung
Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1393-1404

focus on making good solutions to recurring problems accessible and comprehensible and thus, prevent the applier from ‘reinventing the wheel’. Capturing multidisciplinary knowledge related to certain design problems, patterns may decrease cost and effort when solving the problem. Thus, they ensure that already identified requirements and good designs are used (avoidance of potential mistakes), reveal knowledge in an understandable form for all stakeholders and provide a common vocabulary for experts to communicate about design questions. Furthermore, they lead to more transparency of the problem context, as they entail information about structural relations and dependencies that would otherwise probably be neglected [2, 6–8]. Therefore, the use of patterns, has a high impact on the performance of systems, services and organizations, leading to economic benefits and increased competitiveness. Prominent examples of patterns are the Design Patterns by the Gang of Four, which contains patterns about good software engineering [8].

However, although a lot of pattern repositories and catalogues exist and patterns are used in systems engineering practice, the question of how patterns are developed to reach certain design goals (such as quality, applicability or utility) remains unclear. Furthermore, since pattern development is a practical problem and is mostly focused by practitioners or practitioner-oriented communities, existing approaches still lack a scientific foundation. However, we suggest that such a methodological underpinning is necessary in order to have patterns becoming an acknowledged topic for IS and adjacent research fields. We believe that pattern development, underpinned by methodology, can have a strong impact and offer new solutions to existing design problems, as well as increasing their added value through best possible development. From a practical point of view, a clear methodological approach on how to develop patterns is suggested to resolve issues concerning pattern structure, quality and utility.

To solve these problems, we are undertaking a systematic literature review to identify, cluster and structure pattern development approaches. Thus, our work is devoted to structure the field and can be considered a first step towards approaching pattern development from a scientific base.

Our paper is structured as follows. Section 2 gives deeper insights about patterns and pattern development. Our research methodology is described in section 3. The findings of the literature review are given in section 4 and further discussed in section 5. We conclude with our contributions, limitations and future outlook in section 5.

2 Background

2.1 Patterns in IS

The idea to establish and apply patterns originally stems from the field of architecture. Alexander [9] defines patterns as specifications of recurring problems and proven solutions to these problems. Hence, a pattern comprises applied and established knowledge on how to solve a certain design problem. In the mid 1990’s, the pattern concept was first introduced for system development by Gamma et al. [10] who described design patterns for object-oriented software development. Since then, the acceptance and use of patterns has increased tremendously. Nowadays, a huge amount of patterns can be found for different fields of IS, such as human-computer interaction, security, requirements engineering and enterprise integration. Due to the wide circulation of pattern publications, the pattern community is continuously becoming larger. Today, many workshops on scientific conferences (such as RePa[1]) or even entire conferences (e.g., PLoP[2] conferences) are devoted to pattern development and dissemination.

Across all disciplines, the main goal of patterns is reuse of established knowledge for specific problem contexts [11]. Therefore, patterns contain templates for problem specification and other relevant information in tabular form, including a unique pattern name, meta data (e.g., author), conditions of pattern applicability and relations to other patterns [7, 12]. Thus, patterns should contain at least five central elements: a description of *problems* and *forces* to describe the problem and why it is difficult to solve (problem domain), a description of the *solution* and possible *consequences* to demonstrate what happens when the solution is applied (solution domain) and a description of the *context* in which the pattern is applicable [13].

2.2 Pattern Development

Although nowadays many patterns exist for various application domains, literature about how to systematically develop them is scarce. Wellhausen and Fießer [13] introduce the topic by explaining what patterns are and how to identify them with regard to IS disciplines from a practitioners point of view. Furthermore, Withall differentiates between a systematic (inductive) and a theory-based or opportunistic (deductive) approach for pattern development [7]. According to his work, following a systematic approach, is appropriate in case enough satisfactory solutions exist for a problem. In this case, these solutions (e.g., design guidelines) are collected, aggregated and generalized in a pattern format. The opportunistic approach is applicable if no or few solutions exist. Pattern development is then based on relevant experiences and theoretical insights (e.g., empirical data). However, the process itself still lacks a theoretical and methodological foundation which is important for pattern development in a scientifically acknowledgeable manner. Thus, Petter et al. consider the 'scientific' way of pattern development a design science approach, in which the pattern becomes the design artifact [14]. A design artifact is a solution to a problem, which creates value while being used. To ensure the quality of an artifact and therefore its value, it iterates within three different cycles until it creates a satisfactory value. These cycles can be assigned to pattern development for developing satisfactory, value adding, patterns. The Design science research (DSR) cycles can be described as follows [15]. The *relevance cycle* is about identifying the problem

[1] International Workshop on Requirements Patterns held on the International Requirements Engineering Conference. http://re16.org/pages/conference/workshops/#WS08 (accessed October 12, 2016)

[2] Pattern Languages of Programs conferences. http://hillside.net/conferences (accessed October 12, 2016)

domain, eliciting requirements and, after the solution is built, field-testing the artifact. In the *rigor cycle* a common knowledge base is created and other relevant research (e.g., other pattern repositories) is screened to derive design considerations. The goal of the *design cycle* is to build the artifact and validate it against preliminarily defined evaluation criteria. In this work, we follow this logic, since other research present DSR as a suitable way for pattern development [16]. Hence, the conceptual model for our literature review is divided into relevance (identify problems), rigor (include existing knowledge), design (build artifacts) and evaluation (assure value).

3 Methodology

To identify approaches for pattern development, we conduct a systematic literature review [17, 18]. According to the taxonomy given by Cooper, our review is characterized as follows [19]: We focus on approaches for but not the results of pattern development (i.e., the actual patterns). We thereby understand an approach as a technique or type of (collaborative) work, that is used for developing patterns. The goal is to identify pattern development approaches and to integrate them within the DSR framework. The findings will be clustered and organized on a conceptual level. Our target audience are scholars and practitioners. The coverage of this research is representative and should identify common approaches.

The search for relevant articles was conducted in February 2016. For identifying relevant databases and search phrases, a brief research was done based on the aforementioned foundational work including a short forward search. The following databases were chosen, as they cover the PloP conferences, fit the targeted domain of systems engineering and Requirements Engineering: ACM, SpringerLink, IEEE Xplore und AISeL. The predominant phrases in the context of pattern development are: *pattern mining* [7, 20–22], *pattern writing* [7, 22–24] or *pattern authoring* [24, 25]. Unsuitable phrases were *pattern detection* [27, 28] and *pattern discovery* [29], since they are predominantly used in Data Mining or technical pattern detection and hence were excluded from our search. The search was conducted in title, abstract and keywords, which led to 1156 hits. In a first step title, abstract and keywords were screened and unsuitable articles disregarded. In the second step the remaining full articles were read thoroughly to further sort out non-relevant articles. In general, articles were disregarded if they were purely technological, did not mention any approach or were not available in German or English. In the end, a total number of 28 articles were selected for further analysis.

The initial list of possible approaches found in the relevant articles contained 76 items. A four phase process was used to cluster these items into useful and relevant approaches. First, all items were analyzed regarding their abstraction level, which resulted in disregarding very abstract approaches (e.g., Grounded Theory). Second, the meaning of each remaining approach was summarized in a short description- Approaches with similar meaning were clustered in a third step. Fourth, all approaches that did not have at least three hits were disregarded. Since our way of clustering approaches is comparable to coding in qualitative content analysis (i.e., interpretative nature), two researchers clustered the findings independently to establish inter-coder reliability [30]. The clustered approaches were finally analyzed regarding their fit into DSR elements as mentioned above.

4 Approaches for Pattern Development

The findings of the literature review are presented in Table 1. In general, it can be observed that inductive approaches (i.e., observing the specifics of a context and moving towards generalization) are more commonly used to identify and design possible patterns, whereas deductive approaches (i.e., starting from generalized insights and moving towards specifics of a context) are used to structure and validate them in different practical contexts [21, 23]. For example, Schadewitz and Jachna used an inductive approach to find reoccurring issues in student teams' interactions and communications as well as a deductive approach to discover possible connections between these issues and to compare them within different contexts [26]. Another general finding is the importance of collaboration and communication. It urges the meaning of the community and inclusion of interdisciplinary stakeholders for actively creating and sharing comprehensible patterns [24, 25].

Table 1. Pattern Development approaches

Approach	Hits	Relevance	Rigor	Design	Evaluate	Sources
development workshop	15	●	○	●	○	[21, 23–25, 31–41]
enhancement workshop	15	○	○	●	●	[21, 23–25, 31–41]
guided development	15	○	○	●	○	[7, 24, 25, 31–33, 36, 42–46]
shepherding	9	○	○	●	○	[21, 23–25, 32–34, 36, 41]
expert interview	8	●	○	○	●	[26, 32, 33, 35, 39, 41, 43, 47]
observation	8	●	○	○	○	[21, 23, 26, 31–33, 40, 43]
open channel	6	●	○	○	●	[21, 24, 25, 36, 41, 43, 44]
literature review	4	○	●	○	○	[32, 32, 36, 42, 48, 49]
collaborative learning and development	3	●	○	●	○	[33, 35, 39]
pattern mapping	3	●	○	●	○	[21, 26, 32]
pattern writing	3	●	○	●	○	[23, 31, 35]

○ = not applicable / unknown; ● = applicable

Shepherding and workshops are state-of-the-art pattern development approaches and are frequently used on pattern conferences [25]. They are both sharing a highly interactive character. A workshop is a cooperation of a group of people, while shepherding is a cooperation between a shepherd (experienced pattern author) and a sheep (pattern author). A general benefit of a workshop is the attendance of multiple participants which leads to a harmonization of different views [32]. The workshop can target a wide range of goals, from creatively coming up with new pattern ideas to finalizing one definite pattern. Accordingly, there is no generic workshop approach. We suggest two different approaches. First, the pattern development workshop, which is used to mine (collect, categorize and summarize) patterns [23]. For example, Iacob uses a set of workshops to confront designers with a set of problems. The design process is analyzed and transformed into design issues, which are then counted. Their degree of recurrence is calculated to identify pattern candidates [21]. The output of the development workshops are pattern candidates or patterns. Hence, development workshops can be used to find relevant issues and design pattern artifacts. Second, the enhancement workshop, which is about giving and getting feedback on a pattern (evaluation) and using this to improve the pattern until its final form. To continue with the example from Schadewitz and Jachna, their design issues were evaluated in workshops with novices and experts. The insights from those workshops were used for further development. Experts can actively interact with design suggestions, while novices can offer great insights about the actual quality of the pattern [41]. One technique, "fly on the wall", is described very often and helps exemplifying the intention of the enhancement workshop: the pattern author is listening to a group of participants which discusses the pattern he developed without being allowed to intervene or rectify. The goal of this approach is to evaluate the understandability of patterns as stand-alone, self-explaining artifacts [6]. Therefore, the enhancement workshop can be used to design and evaluate patterns.

Shepherding is much more focused on pattern specific knowledge and the artifact itself. It usually takes place over several months and is an intense interaction between the shepherd and the sheep. This makes it a very valuable approach for the design phase. In practice, it is often used in advance of pattern conferences, which offer a wide range of approaches to support pattern development. They contain multiple approaches like workshops, discussions, presentations and shepherding.

Collaborative learning and development, is an effective approach to start pattern mining, as it teaches domain experts and novices about patterns and how to develop them. While understanding the nature of patterns and their use, participants start recognizing and can support each other in developing first patterns. This approach raises pattern awareness, which helps finding relevant issues and supports the design of first patterns.

Guided development is mainly used to support designing the actual pattern as it informs about possible formats (even offers tool supported templates), steps to do while writing a pattern or best practices with useful recommendations. It makes particular pattern development knowledge available to be used by authors. The approach comprises elements such as step-by-step guides, pattern templates and checklists.

Expert interviews are interactive meetings of two people or a small group to gather each subject's personal views and experiences within their context. The interviewees are either experts of a certain domain and thus act as source for issues within their expertise (relevance) or future pattern appliers who give feedback on pattern designs (evaluation).

A not necessarily interactive approach for investigating people and/or technology within their realworld environment, is observation. It was also used for first patterns by Alexander [50] and is highly applicable in identifying relevant issues for pattern development. Especially for non-communicable complex situations and contexts this approach is highly effective.

The open channel approach implies communication between all stakeholders of a system. A continuous discussion between stakeholders leads to two main benefits: Already discovered patterns are being continuously evaluated and new issues for pattern development might arise within a simple discussion. Therefore, the open channel approach is inherent in most of the presented approaches. However, its core intention is to focus on being open for discussion at all other phases as well. Fehling et. al. posit that patterns evolve continuously through discussion in a community [44].

Pattern mapping is used to structure codes, elements or domains of patterns with regards to their relationships. This approach supports the design of patterns as it sheds light on interdependencies between different patterns. Additionally, it can be used to identify relevant issues or missing solutions for new patterns as it reflects the structure of the underlying problem domain. [32].

The literature review approach is a systematic research of existing knowledge about existing solutions, a theoretical base or adjacent patterns. A literature review can access a multitude of resources from wikis, scientific and practitioner-oriented outlets or pattern repositories [32]. Pattern repositories, as an important resource for pattern development, comprise different patterns for certain problem domains or applications [48]. However, due to the vast amount of potential sources, a pattern-oriented literature review may be a difficult task.

Pattern writing is the actual process of creating the pattern artifact, but it is also an approach for identifying relevant patterns as it reveals new facts by adding additional experiences or asking the authors to express more of their implicit knowledge [23].

5 Discussion

It can be observed that most of the identified approaches are applicable in the design cycle. Thus, it may be inferred that this is the core to the very practitioner-oriented and less scientific state-of-the-art of pattern development. However, we second the opinion that this phase is the most crucial part of a pattern development process. However, high number of approaches can be applied in the relevance cycle as well, which again highlights that pattern development is practically-driven process. Not many approaches were found applicable for the rigor or evaluation phase. Although a vast amount of pattern repositories exist that can potentially be used in the rigor cycle, the search for patterns relevant to a specific development problem is often a complex endeavor. This is mainly caused by a lack of structuration as well as nonexistent formatting standards for patterns and pattern repositories. Furthermore, validation of patterns as usually conducted in the design-evaluation phase of DSR has yet widely been neglected. This issue is also manifested in the vague definition of when a recurring solution can be considered a pattern: the pattern community calls for at least three good implementations, without specifying what 'good' means. Hence, based on our findings, we call for rigor in (1) screening existing relevant knowledge that may be useful for pattern design and (2) validate patterns and evaluate them in the problem context. We believe that both are necessary to give pattern development a stronger scientific impact besides its practical relevance.

6 Conclusion and Research Outlook

Our research embodies a first step towards a clearer understanding of pattern development approaches from a design science perspective. To the best of our knowledge, we are the first to structure existing approaches for pattern development by conducting a systematic literature review. With this contribution we aim to make the following contributions: From a scholarly perspective, we made steps towards an operationalization of DSR for pattern development, which leads to the creation of both prescriptive and descriptive design knowledge [51, 52]. From a practical viewpoint, pattern authors can use our findings to enhance their own pattern development. Our findings serve as a guide through necessary phases of pattern development and highlight applicable approaches for each phase. Furthermore, the description provided for each approach supports pattern authors in choosing the best fit for their context. This may lead to patterns with improved quality and more efficient and effective pattern development.

However, since our research is only a first step towards structuring the field it does not come without limitations. Although we are sure in having covered the most relevant approaches with our orientation towards the pattern community, literature search should be extended to more scientific and pattern-related outlets and databases in future research in order to reveal complementary approaches and enhance the presented findings. Furthermore, future research may build up on our results by creating (reference) processes for pattern development. Finally, evaluation of our approach in specific problem contexts is necessary to prove our results in field.

References

1. Pohl, K.: Requirements Engineering. Grundlagen, Prinzipien, Techniken. Dpunkt-Verl., Heidelberg (2008)
2. Towards Interdisciplinary Design Patterns for Ubiquitous Computing Applications. Kassel University Press, Kassel (2014)
3. Hoffmann, A., Hoffmann, H. & Söllner, M.: Fostering Initial Trust in Applications – Developing and Evaluating Requirement Patterns for Application Websites. Utrecht, Netherlands (2013)

4. Hoffmann, A., Söllner, M., Hoffmann, H., Leimeister, J.M.: Requirement Patterns to Support Socio-Technical System Design. In: David, K., Geihs, K., Leimeister, J.M., Roßnagel, A., Schmidt, L., Stumme, G., Wacker, A. (eds.) Socio-technical Design of Ubiquitous Computing Systems, pp. 191–209. Springer International Publishing, Cham (2014)
5. Gebauer, L., Kroschwald, S. & Wicker, M.: Anforderungsmuster zur Förderung der Rechtmäßigkeit und Rechtsverträglichkeit von Cloud Computing-Diensten. Kassel University Press, Kassel (2015)
6. Kohls, C.: The theories of design patterns and their practical implications exemplified for e-learning patterns. Ingolstadt (2014)
7. Withall, S.: Software requirement patterns. Microsoft Press, Redmond, Wash. (2007)
8. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design patterns. Elements of resuable object-oriented software. Pearson education limited, England (1995)
9. Alexander, C.: The timeless way of building. Oxford University Press, New York (1979)
10. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns - Elements of Reusable Object-Oriented Software. Addison-Wesley, Reading (1994)
11. Franch, X., Palomares, C., Quer, C., Renault, S., Lazzer, F. de: A Metamodel for Software Requirement Patterns. In: Wieringa, R., Persson, A. (eds.) Requirements Engineering: Foundation for Software Quality, 6182, pp. 85–90. Springer Berlin Heidelberg (2010)
12. Durán Toro, A., Bernárdez Jiménez, B., Ruiz Cortés, A., Toro Bonilla, M.: A Requirements Elicitation Approach Based in Templates and Patterns. Workshop em Engenharia de Requisitos 1999 (1999)
13. Wellhausen, T. and Fießer, A.: How to write a pattern? A rough guide for first-time pattern authors, http://europlop.net/sites/default/files/files/0_How%20to%20write%20a%20pattern-2011-11-30_linked.pdf
14. Petter, S., Khazanchi, D., Murphy, J.D.: A Design Science Based Evaluation Framework for Patterns. SIGMIS Database 41, 9–26 (2010)
15. Hevner, A.R.: A three cycle view of Design Science Research. Scandinavian Journal of Information Systems 19, 87–92 (2007)
16. Hoffmann, A.: Anforderungsmuster zur Spezifikation soziotechnischer Systeme. Standardisierte Anforderungen der Vertrauenwürdigkeit und Rechtsverträglichkeit. Kassel University Press, Kassel, Germany (2014)
17. Vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A. (eds.): Reconstructing the giant: On the importance of rigour in documenting the literature search process (2009)
18. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly (2002)
19. Cooper, H.M.: Organizing knowledge syntheses. A taxonomy of literature reviews. Knowledge in Society 1, 104–126 (1988)
20. Kohls, C.: The structure of patterns. In: Kohls, C. (ed.) the 17th Conference, pp. 1–10 (2010)

21. Iacob, C.: A design pattern mining method for interaction design. In: Paternò, F., Luyten, K., Maurer, F. (eds.) the 3rd ACM SIGCHI symposium, p. 217 (2011)
22. Hillside Group: The Hillside Group - A group dedicated to design patterns. Home of the patterns library., http://hillside.net/
23. Kohls, C., Panke, S.: Is that true…? In: Wirfs-Brock, R. (ed.) the 16th Conference, p. 1 (2009)
24. Wellhausen, T., Fiesser, A.: How to write a pattern? In: Avgeriou, P., Fiesser, A. (eds.) the 16th European Conference, pp. 1–9 (2011)
25. Fehling, C., Barzen, J., Breitenbücher, U., Leymann, F.: A process for pattern identification, authoring, and application. In: Eloranta, V.-P., van Heesch, U. (eds.) the 19th European Conference, pp. 1–9 (2014)
26. Schadewitz, N., Jachna, T.: Introducing New Methodologies for Identifying Design Patterns for Internationalization and Localization. In: Aykin, N. (ed.) Usability and Internationalization. Global and Local User Interfaces. Second International Conference on Usability and Internationalization, UI-HCII 2007, Held as Part of HCI International 2007, Beijing, China, July 22-27, 2007, Proceedings, Part II, 4560, pp. 228–237. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg (2007)
27. Yu, D., Zhang, Y., Ge, J., Wu, W.: From Sub-patterns to Patterns: An Approach to the Detection of Structural Design Pattern Instances by Subgraph Mining and Merging. In: 2013 IEEE 37th Annual Computer Software and Applications Conference (COMPSAC), pp. 579–588 (2013)
28. Tsantalis, N., Chatzigeorgiou, A., Stephanides, G., Halkidis, S.: Design Pattern Detection Using Similarity Scoring. IIEEE Trans. Software Eng. 32, 896–909 (2006)
29. Gupta, M., Pande, A., Rao, R.S., Tripathi, A.K.: Design Pattern Detection by normalized cross correlation. In: 2010 International Conference on Methods and Models in Computer Science (ICM2CS 2010), pp. 81–84
30. Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken. Beltz, Weinheim u.a. (2010)
31. Meszaros, G., Doble, J.: A pattern language for pattern writing. In: Martin, R.C., Riehle, D., Buschmann, F. (eds.) Pattern languages of program design 3. [papers to PLoP '96 and EuroPLoP '96]. Addison-Wesley, Reading, Mass. (1998)
32. Lotz, N., Law, E.L.-C., Nguyen-Ngoc, A.V.: A process model for developing learning design patterns with international scope. Education Tech Research Dev 62, 293–314 (2014)
33. Linden, T., Cybulski, J.: Refining the process of sharing problem-solving experience across domain: A hermeneutic study. AMCIS 2006 Proceedings, 433 (2006)
34. Rising, L., Rehmer, K.: Patterns for sustainable development. In: Kohls, C. (ed.) the 17th Conference, pp. 1–11 (2010)
35. Wesson, J., Cowley, L.: UI Design Patterns: From Theory to Practice. In: Seffah, A., Desmarais, M.C., Gulliksen, J. (eds.) Human-Centered Software Engineering - Integrating Usability in the Software Development Lifecycle, 8, pp. 331–351. Springer, Dordrecht (2005)

36. Schümmer, T., Haake, J.M.: Shaping Collaborative Work with Proto-patterns. In: Pipek, V., Rosson, M.B., Ruyter, B. de, Wulf, V. (eds.) End-user development. 2nd international symposium, IS-EUD 2009, Siegen, Germany, March 2 - 4, 2009 ; proceedings, 5435, pp. 166–185. Springer, Berlin (2009)
37. Köppe, C.: Using pattern mining for competency-focused education. In: van der Veer, G., Sikorski, M., Sloep, P., van Eekelen, M. (eds.) Second Computer Science Education Research Conference, pp. 23–26 (2012)
38. Deng, J., Kemp, E., Todd, E.G.: Focussing on a standard pattern form. In: Billinghurst, M. (ed.) the 6th ACM SIGCHI New Zealand chapter's international conference, pp. 83–90 (2006)
39. Akado, Y., Kogure, S., Sasabe, A., Hong, J.-H., Saruwatari, K., Iba, T.: Five Patterns for Designing Pattern Mining Workshops. In: EuroPLoP '15, July 08-12, 2015, Kaufbeuren, Germany (2015)
40. Hagge, L., Houdek, F., Lappe, K., Paech, B.: Using Patterns for Sharing Requirements Engineering Process Rationales. In: Dutoit, A.H., McCall, R., Mistrík, I., Paech, B. (eds.) Rationale Management in Software Engineering, pp. 409–427. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg (2006)
41. Hentrich, C., Zdun, U., Hlupic, V., Dotsika, F.: An approach for pattern mining through grounded theory techniques and its applications to process-driven SOA patterns. In: Kohls, C., van Heesch, U. (eds.) the 18th European Conference, pp. 1–16 (2015)
42. Diaz, P., Acuña, P., Aedo, I., Malizia, A.: A Design Patterns Catalog for Web-Based Emergency Management Systems. In: D'Atri, A., Marco, M. de, Braccini, A.M., Cabiddu, F. (eds.) Management of the Interconnected World. ItAIS: The Italian Association for Information Systems, pp. 387–394. Springer-Verlag Berlin Heidelberg, Heidelberg (2010)
43. Välimäki, A., Kääriäinen, J.: Requirements Management Practices as Patterns for Distributed Product Management. In: Münch, J., Abrahamsson, P. (eds.) Product-focused software process improvement. 8th international conference, PROFES 2007, Riga, Latvia, July 2 - 4, 2007 ; proceedings, 4589, pp. 188–200. Springer, Berlin (2007)
44. Fehling, C., Ewald, T., Leymann, F., Pauly, M., Rutschlin, J., Schumm, D.: Capturing Cloud Computing Knowledge and Experience in Patterns. In: 2012 IEEE 5th International Conference on Cloud Computing (CLOUD), pp. 726–733 (2012)
45. Borchers, J.: The Aachen Media Space: Design Patterns for Augmented Work Environments. In: Lahlou, S. (ed.) Designing user friendly augmented work environments. From meeting rooms to digital collaborative spaces, pp. 261–312. Springer, London, New York (2009)
46. Gholami, M.F., Jamshidi, P., Shams, F.: A Procedure for Extracting Software Development Process Patterns. In: 2010 European Modelling Symposium (EMS), pp. 75–83 (2010)
47. Benedicenti, L., Succi, G., Vernazza, T.: From process modeling to domain modeling. SIGAPP Appl. Comput. Rev. 5, 28–32 (1997)

48. Franch, X., Renault, S., Mendez-Bonilla, O., Quer, C.: PABRE: Pattern-based Requirements Elicitation. In: Third International Conference on Research Challenges in Information Science (RCIS), pp. 81–92 (2009)
49. Welicki, L., Manuel, J., Lovelle, C., Aguilar, L.J.: Patterns meta-specification and cataloging. In: Tarr, P., Cook, W.R. (eds.) Companion to the 21st ACM SIGPLAN conference, p. 679 (2006)
50. Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M.: A pattern language. Towns, buildings, construction. Oxford Univ. Press, New York, NY (1977)
51. Gregor, S.: The Nature of Theory in Information Systems. MIS Quarterly 30, 611–642 (2006)
52. Chandra, L., Seidel, S., Gregor, S.: Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions. In: 2015 48th Hawaii International Conference on System Sciences (HICSS), pp. 4039–4048

Text Mining in Ideencommunities: Gestaltungsempfehlungen eines Text Mining Artefakts zur Unterstützung der Ideenevaluation

Thomas Maier[1], Sebastian Gronich[1], Ann-Katrin Rickert[1] and Nikolaus Lipusch[2]

[1] University of Kassel, Kassel, Germany
thomas_maier10@web.de, Basti.Oliver@gmx.de, anka.rickert@googlemail.com

[2]University of Kassel, Information Systems, Kassel, Germany
nikolaus.lipusch@uni-kassel.de

Abstract. Ideencommunities haben sich als effektives Tool zur Erhöhung der Innovationskraft von Unternehmen etabliert. Ein Aspekt, dem bislang wenig Aufmerksamkeit geschenkt wurde, ist die Ideenevaluation in Ideencommunities. Ein Problem im Rahmen der Evaluation ist die Vielzahl an Ideen, die in solchen Communities entsteht und Unternehmen zunehmend vor neue Herausforderungen stellt. So verfügen Unternehmen in der Regel nur über eine begrenzte Kapazität, um das vollständige Potenzial solcher Ideen nutzbar zu machen. Eine Lösung für dieses Problem bieten sogenannte Text Mining Verfahren, die in Verbindung mit neuen Informations- und Kommunikationstechnologien Unternehmen dabei unterstützen können, Ideen effizienter auszuwerten sowie neue Datenzusammenhänge und somit wertvolle Information zu erkennen. Vor diesem Hintergrund untersucht der Beitrag welche Phasen der Ideenevaluierung am effizientesten durch ein Text Mining Tool unterstützt werden können und leitet daraus konkrete Gestaltungsempfehlungen für ein Text Mining Tool für Ideencommunities ab. Der Beitrag umfasst dabei konkrete Gestaltungsempfehlungen, die Praktiker bei der Umsetzung eines solchen Tools unterstützen sollen. Der theoretische Beitrag liegt in der Erweiterung des Erkenntnisstandes hinsichtlich neuer Methoden zur Unterstützung der effizienten Auswertung von Ideen im Crowdsourcing

Keywords: Ideencommunities, Absorpative Capacity, Text Mining, Design Science Research

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Maier, T.; Gronich, S.; Rickert, A.-K.; Lipusch, N. (2017): Text Mining in Ideencommunities: Gestaltungsempfehlungen eines Text Mining Artefakts zur Unterstützung der Ideenevaluation, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1405-1416

1 Einleitung

In jüngerer Vergangenheit hat Crowdsourcing immer mehr an Popularität gewonnen [1]. Eine spezielle Form des Crowdsourcing, die von Unternehmen zunehmend aktiv eingesetzt wird, sind sogenannte Ideencommunities. Es handelt sich dabei zumeist um Plattformen, die es Unternehmen ermöglichen bestimmte Probleme zu posten und in Zusammenarbeit mit der Crowd Lösungen in Form von Ideen zu generieren. Das Ziel von Ideencommunities ist es dabei, von den Erfahrungen und Ideen der Kunden zu profitieren und dadurch die Innovationskraft des Unternehmens zu erhöhen. Die Vorteile von Ideencommunities wurden in der Literatur umfangreich belegt [2]. Das Unternehmen Lego kann hierbei exemplarisch für den Aufbau einer solchen Ideencommunity hervorgehoben werden [3]. Durch das interaktive Kollaborieren und Zusammenwirken der Crowd wird der dem Ansatz zugrunde liegende Kerngedanke des Nutzbarmachens der sogenannten „Weisheit der Vielen" deutlich [4, 5].

Trotz der genannten Vorteile, die Ideencommunities mit sich bringen, ist deren effiziente Nutzung auch mit großen Herausforderungen verbunden. So entsteht durch die rege Teilnahme an Ideencommunities eine Vielzahl an Beiträgen. Das Resultat ist eine große Menge an unstrukturierten Daten. My Starbucks Idea umfasst beispielsweise über mehrere Kategorien hinweg ca. 150.000 Ideenvorschläge, Dell Idea Storm über 25.000 (Stand: September 2016). Die große Menge an textuellen Informationen macht es dabei in der Regel schwierig das volle Potenzial von Ideencommunities zu nutzen [6, 7]. Unternehmen, als Betreiber solcher Ideencommunities, sehen sich dabei mit der Herausforderung konfrontiert, dass sie nur über begrenzte Ressourcen und Kapazitäten verfügen, um diese Daten auszuwerten. Die beschriebene Problematik wird in der Literatur unter dem Begriff Absorptive Capacity thematisiert. Hierunter wird die Fähigkeit verstanden, Wissen aus externen Quellen zu identifizieren, assimilieren und anschließend anzuwenden [8]. Die Mengen an unstrukturierten Daten, die im Rahmen von Ideencommunities generiert werden, stellen eine besondere Herausforderung für die Absorptive Capacity von Unternehmen dar. Eine Möglichkeit, Unternehmen bei der Auswertung von Ideencommunities zu unterstützen, bieten moderne Informations- und Kommunikationstechnologien. Neben der effektiven Einbeziehung von Kunden bergen diese das Potenzial Organisationsroutinen und die Koordinationsaktivitäten von Unternehmen zu verbessern und Kosten zu senken [9, 10].

Ein vielversprechender Ansatz, um Unternehmen bei der IT-gestützten Auswertung von Ideen zu unterstützen, bilden sogenannte Text Mining Anwendungen [41]. Es handelt sich dabei um Tools, die eine Kombination aus computergestützten linguistischen, statistischen und mathematischen Verfahren anwenden, um Wissen aus großen Mengen an unstrukturierten Textdaten heraus zu filtern und neue Zusammenhänge zu erkennen [11]. Obwohl von einem hohen Potenzial von Text Mining Lösungen für Ideencommunities ausgegangen werden kann, blieb eine kombinierte Betrachtung dieser beiden Forschungsgebiete bislang unberücksichtigt [1, 12]. Demnach wurde lediglich partiell abgedeckt, wie sich eine Text Mining Lösung systematisch für Ideencommunities nutzen lässt. Vor diesem Hintergrund ist es Ziel dieses Papers potenzielle Einsatzgebiete und Unterstützungspotenziale eines Text

Mining Tools zur Ideenevaluierung in Ideencommunities zu identifizieren und Unternehmen Gestaltungsempfehlungen für ein Tool zur automatisierten und somit effizienteren Ideenevaluation an die Hand zu geben. Das generelle Vorgehen im Rahmen der Studie orientiert sich dabei am Design Science Research (DSR)-Ansatz nach Vaishnavi und Kuechler [13].

Der sich hieraus erschließende Beitrag bezieht sich auf zwei Bereiche. Zum einen geben wir Praktikern Empfehlungen zum Einsatz und zur Gestaltung eines Text Mining Tool an die Hand, die darauf abzielen den Prozess der Ideenevaluierung in Ideencommunities zu automatisieren und Unternehmen bei der Analyse von Ideen zu unterstützen. Der theoretische Beitrag dieser Arbeit liegt in einer Erweiterung des Erkenntnissstandes zur Ideenauswertung im Crowdsourcing. Das Ziel hierbei ist es die Phase der Ideenauswertung näher zu beleuchten. Speziell die einzelnen Prozessschritte, die mit der Phase der Ideenevaluation einhergehen, werden in der vorliegenden Arbeit betrachtet, da die genaue Untersuchung dieser Thematik in der Literatur bisweilen noch keinen Eingang fand. Darüber hinaus soll gezeigt werden, wie ein Text Mining Tool in den einzelnen Prozessschritten am besten unterstützen kann. Wir setzen hierbei an vorangehenden Arbeiten zum Crowdsourcing an, die zu der Erkenntnis kommen, dass die Phase der Ideenevaluation ein großes Potenzial für die Effizienzsteigerung und die breite Anwendung von Crowdsourcingaktivitäten birgt.

Das folgende Paper gliedert sich wie folgt. Nach der Einführung des theoretischen Hintergrunds wird die zugrunde gelegte Methode in Form des DSR-Ansatzes genauer erläutert. Im Anschluss an die Ermittlung des Einsatzpotenzials eines Text Mining Tools zur Unterstützung der Ideenevaluierung und der konkreten Ableitung von Gestaltungsempfehlungen, welche ein Text Mining Tool erfüllen muss, werden die weiteren geplanten Schritte für die Entwicklung dieses Tools dargelegt. Das Paper schließt mit einer Zusammenfassung und dem erwarteten Beitrag für Praxis und Wissenschaft.

2 Theoretischer Hintergrund

2.1 Absorptive Capacity Theorie

Das Grundkonzept der Absorptive Capacity wird in der Fachliteratur als Fähigkeit von Unternehmen definiert „to recognize the value of new, external information, assimilate it, and apply it to commercial ends“ [8]. Aus der Definition ist abzuleiten, dass sich die Absorptionsfähigkeit in sequentielle Teilfähigkeiten differenzieren lässt [14].

Zunächst beschreibt es die Fähigkeit eines Unternehmens neue, externe Informationen zu identifizieren und dabei den zu schöpfenden, potenziellen Wert zu erkennen. In einem weiteren Schritt müssen im Zuge der Assimilationsfähigkeit die Informationen an die Bedürfnisse adäquat angepasst bzw. an das bestehende Wissen des Unternehmens angeknüpft werden. Abschließend kann das Wissen angewendet bzw. kommerziell eingesetzt werden [8, 15]. Dieses klassische Konzept zur Absorptive Capacity nach Cohen und Levinthal [8] wurde in den darauffolgenden Jahren insbesondere durch die Beiträge von Zahra und George [16] sowie Todorova und Durisin [17] theoretisch diskutiert, konkretisiert und erweitert. Lane et al. [18] verstehen Absorptive Capacity als eines der zentralen Konstrukte, wie Unternehmen oder Organisationseinheiten Wissen einarbeiten, welches sich außerhalb ihrer Grenzen befindet. Im Zuge des Trends Richtung Open Innovation werden Kunden in die Gestaltung von Innovationsprozessen der Unternehmen aktiv integriert [19]. In diesem Zusammenhang werden Ideencommunities als Instrumente eingesetzt, um Ideen bzw. Wissen zu generieren. Die Herausforderung liegt explizit darin, die Unmengen an unstrukturierten Daten den Bedürfnissen eines Unternehmens anzupassen, um anschließend relevantes Wissen absorbieren und anwenden zu können. Zur effizienten Steuerung der Informationen im Unternehmen legen Roberts et al. [9] den Einsatz von Informationstechnologien dar, welche in Form von automatisierten Anwendungstools in den Absorbierungsprozess integriert werden. Das vorliegende Paper setzt sich mit der Verwendung eines Text Mining Tools auseinander, welches unter Berücksichtigung der Effizienzsteigerung im Absorbierungsprozess eingesetzt werden soll.

2.2 Grundlagen zum Crowdsourcing

Der Neologismus aus den beiden Wörtern „outsourcing" und „crowd" wird in Anlehnung an Leimeister [20] als ein Interaktionsprozess definiert. Eine Unternehmung, Organisation, Gruppe oder ein Individuum, auch Crowdsourcer genannt, bietet einer nicht genauer definierten großen Gruppe, den sogenannten Crowdsourcees, die Möglichkeit der Bearbeitung einer Aufgabe über einen offenen Aufruf. Ziel ist es, sich das aggregierte Wissen der Masse zu Nutze zu machen. Die Crowdsourcees, welche sich ebenfalls aus Individuen, formellen sowie informellen Gruppen, Organisationen oder Unternehmen zusammensetzen können, übernehmen anschließend die Aufgabenbearbeitung [20].

Eine besondere Ausprägungsform dieses genannten Ansatzes sind von Unternehmen initiierte und betriebene, IT-basierte Ideencommunities. Sie ermöglichen die Teilnahme von Individuen außerhalb des Unternehmens, insbesondere von Kunden, am unternehmerischen Innovationsprozess [7]. Kunden haben auf Ideencommunities die Möglichkeit, eigene Erfahrungen, Verbesserungsvorschläge sowie Innovationsideen bekanntzugeben, Vorschläge anderer zu bewerten, interaktiv mit Hilfe von Kommentar- und Wiki-Funktionen weiterzuentwickeln und auf diese Weise aktiv zur Produkt- und Serviceentwicklung beizutragen [7].

Die Anwender und Nutzer der Ideencommunities generieren im Rahmen dieses interaktiven Prozesses eine Menge an Ideen, die als textuelle Informationen gespeichert werden. Dieses wird bisher in einem ressourcenintensiven, manuellen Verfahren durch User sowie Experten evaluiert. Die auf eine Person begrenzte und wertvolle Arbeitskraft, die für diesen Auswertungsprozess benötigt wird, steht anschließend nicht mehr für die angemessene Analyse der vielen Beiträge und Ideen zur Verfügung. Darüber hinaus kann die Analyse der Ideen aufgrund der Masse an textueller Information nicht adäquat vorgenommen werden, denn häufig fokussieren die am Analyseprozess beteiligten Experten unbewusst auf einige wenige Ideen, ohne dabei die Gesamtheit der Ideen zu berücksichtigen. Aufgrund dieses zeitintensiven und subjektiven Auswertungsvorgehens gelingt es Unternehmen folglich nicht das gesamte Potenzial der Ideen zu nutzen. Eine Möglichkeit den zuvor genannten Problemen zu begegnen und dabei die Absorbierungsfähigkeit von Unternehmen zu unterstützen, bildet der Einsatz eines automatisierten Text Mining Tool.

2.3 Grundlagen zum Text Mining

Der Begriff Text Mining bezeichnet einen Ansatz, welcher sich aus verschiedenen Verfahren zusammensetzt. Durch die Anwendung unterschiedlicher Techniken werden neue Informationen aus unstrukturierten Texten generiert und diese automatisch extrahiert [21]. Text Mining wird dem Bereich des Data Mining zugeordnet, mit dem Unterschied, dass Informationen im Gegensatz zum Data Mining aus unstrukturierten Daten gewonnen werden [21, 22]. In der Literatur wird überwiegend ein Prozessverlauf thematisiert, welcher das Text Mining Verfahren in zwei Phasen aufteilt [23, 24].

Die erste Phase stellt hierbei die Datenvorbereitungsphase dar. Der unstrukturiert vorliegende Text muss in ein Format transferiert werden, aus dem ein Text Mining Algorithmus Informationen gewinnen kann. Dieser Vorbereitungsprozess bedient sich dabei computerlinguistischer Methoden in Form der morphologischen, syntaktischen und semantischen Analyse [23]. Im Bereich der morphologischen Analyse wird durch die sogenannte Tokenisierung (Entfernen von Satz- und Leerzeichen; Generierung von Tokens), das Filtern (Stoppwortentfernung) sowie das Stemming (Rückführung der Wörter zu Wortstämmen) der Text für den weiteren Verlauf vorbereitet [24, 25]. Ferner behandelt die syntaktische Analyse den Textinhalt und versucht durch Worterkennungen mit Hilfe syntaktischer Regeln einzelne Phasen in strukturierte Syntaxbäume zu bringen. Darüber hinaus werden hierdurch das Verhältnis sowie die gegenseitige Modifikation einzelner Phrasen erkannt [26]. Die semantische Analyse hingegen behandelt Mehrdeutigkeiten der einzelnen Wörter und deren Bezug zum Gesamtkontext [27]. Das Ziel dieser ersten Phase ist demnach die Komplexitätsreduktion, um einen bereinigten und nur aus potenziell interessanten Wörtern bestehenden Text zu erhalten, der in einem nächsten Schritt bearbeitet werden kann [28].

In der zweiten Phase ist es nun durch das Nutzen von Mining-Techniken möglich, wichtige Informationen aus dem überarbeiteten Text zu gewinnen. Daten können hierbei vordefinierten Klassen durch Verfahren wie Naïve Bayes, Logistic regression, Entscheidungsbäume, Neuronale Netzwerke, Support Vector Machines, MARSplines, k-nearest neighbors zugeordnet werden [29, 30, 31, 32, 33]. Clustering-Methoden wie das k-means Clustering [25] können sowohl einzeln, als auch in Kombination mit anderen Methoden angewandt werden, um Gruppen mit denselben Inhalten zu bilden.

Abschließend wurde herausgearbeitet, wie sich Text Mining Methoden auf den Kontext von Ideencommunities anwenden lassen, um aus einer unstrukturierten Menge an Daten wichtige Informationen zu extrahieren, diese in einem Prozess für die weiteren Bearbeitungsschritte zu selektieren, um schlussendlich einen Beitrag für eine effiziente Ideenevaluation leisten zu können.

3 Methode

3.1 Design Science Research

Das vorliegende Paper bezieht sich auf den DSR-Ansatz nach Vaishnavi und Kuechler [13]. Dieser Ansatz konzentriert sich auf die Konstruktion sowie Evaluation von sogenannten Artefakten, welche zur Lösung eines vorab definierten Problems beitragen [35, 36]. Artefakte können hierbei mehrere Ausprägungen annehmen, wie beispielsweise Konstrukte, Modelle, Methoden, Instanziierungen oder wie unserem Fall Design-Guidelines [37].

Das Ziel dieses Papers ist es, Einsatzgebiete und Unterstützungspotenziale eines Text Mining Tools im Rahmen der Ideenevaluierung zu identifizieren, sowie entsprechende Gestaltungsempfehlungen in Form von Design-Guidelines für Unternehmen abzuleiten. Die Design-Guidelines dienen als Basis für die Gestaltung eines Text Mining Tools zur Unterstützung der Ideenanalyse in Ideencommunities. Das Tool soll dabei konzipiert werden um eine große Menge an unstrukturiert vorliegenden Textdaten automatisiert aufzubereiten und auszuwerten. Es dient dabei, Experten in der Ideenevaluation zu unterstützen, indem es eine effektivere Nutzung des Potenzials vorhandener Daten erlaubt.

Vaishnavi und Kuechler [13] untergliedern den DSR-Ansatz in fünf Phasen: Awareness of Problem, Suggestion, Development, Evaluation und Conclusion. In unserem Fall wurden die ersten drei Phasen konkret bearbeitet.

In der ersten Phase, *Awareness of Problem*, wurde zunächst der Prozess der Ideenevaluierung in Ideencommunities mit Hilfe einer systematischen Inhaltsanalyse abgeleitet, um anschließend Probleme in den einzelnen Prozessphasen identifizieren zu können. Insgesamt wurden im Rahmen der Analyse die folgenden fünf Ideencommunities untersucht: My Starbucks Idea, SAPiens, Dell Ideastorm, LEGO Ideas und Tchibo Ideas. Bei den Betreibern der Plattformen handelt es sich um Unternehmen aus unterschiedlichen Branchen, die mit ihren Ideencommunities primär Endnutzer (B2C-Bereich) ansprechen. Darüber hinaus ähneln sich die betrachteten Communities stark im Hinblick auf ihren Aufbau und ihre Struktur. So erlauben es alle Communities Probleme zu posten und Ideen durch User generieren und bewerten zu lassen. Die Analyse folgte einem iterativen Vorgehen, indem 3 Autoren die oben genannten Plattformen unabhängig voneinander untersuchten und anschließend Probleme im Hinblick auf die Ideenevaluation identifizierten.

In der zweiten Phase, *Suggestion*, wurden im Zuge von Interviews mit drei Crowdsourcing Experten potenzielle Einsatzgebiete von Text Mining im Ideenevaluierungsprozess erarbeitet. Bei den Crowdsourcing Experten handelte es sich um Spezialisten mit umfassender Erfahrung hinsichtlich der Durchführung von Crowdsourcingprojekten sowie um Personen, die sich umfassend mit dem Thema aus praktischer und wissenschaftlicher Sicht beschäftigt haben. Als Grundlage für die Interviews dienten der zuvor abgeleitete Prozess und die identifizierten Probleme.

In der dritten Phase, *Development,* wurden auf Basis der zuvor identifizierten Probleme konkrete Gestaltungsempfehlungen in Form von Design-Guidelines abgeleitet. Zur konkreten Ausgestaltung des Text Mining Tools wurden Interviews mit drei Text Mining Experten geführt. Hierbei handelt es sich um Spezialisten auf dem Gebiet der Entwicklung und Implementierung von Text Mining Lösungen. Ein Interview wurde mit dem Geschäftsführer eines Unternehmens geführt, welches sich auf Text Mining Lösungen spezialisiert hat. Die beiden weiteren Interviewpartner verfügen über langjährige Erfahrung in der Implementierung von Text Mining Lösungen und mit der Betreuung von Text Mining Projekten.

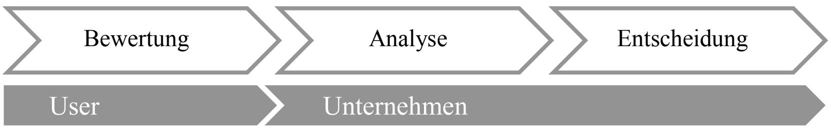
Die weiterführenden Prozessphasen, *Evaluation* und *Conclusion*, werden in Kapitel 5 betrachtet.

3.2 Vorläufige Ergebnisse (Anforderungen an ein Text Mining Tool)

In diesem Paper können wir bereits vorläufige Ergebnisse präsentieren, welche die Ziele und damit einhergehenden Anforderungen eines Text Mining Tools in Bezug auf Ideencommunities kennzeichnen.

In dem ersten Prozessschritt *Awareness of Problem* wurde zunächst, auf Grundlage einer systematischen Inhaltsanalyse von mehreren Ideencommunities, der Prozess zur Ideenevaluierung abgeleitet (siehe Abbildung 1) und damit einhergehend potenzielle Problemfelder beleuchtet.

Abbildung 1. Bisheriger Prozess der Ideenevaluation



Wie in Abbildung 1 ersichtlich gliedert sich der Prozess der Ideenevaluierung in drei Phasen. In Phase eins, *Bewertung*, erfolgt eine initiale Bewertung der Ideen durch die User. Dies geschieht über das Verfassen von Kommentaren oder durch die Anwendung von Bewertungsmechanismen wie Ratingskalen. In Phase zwei, *Analyse*, werden die Ideen durch das Unternehmen analysiert. Diese Phase ist dadurch gekennzeichnet, dass Unternehmen eine Vielzahl an Beiträgen und Kommentaren analysieren sowie bewerten müssen. In Phase drei, *Entscheidung,* entscheidet das Unternehmen, wie und mit wem die Ideen weiterentwickelt bzw. welche Ideen weiterverfolgt werden. Wie aus der obigen Abbildung ersichtlich werden Beiträge und die daraus resultierende Menge an unstrukturierten Textdaten im bisherigen Prozesses der Ideenevaluation durch User und durch Experten evaluiert. Dem Prozess liegt dabei eine stark subjektive Betrachtung- und Vorgehensweise zu Grunde.

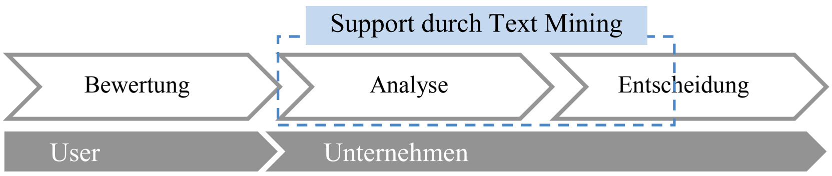
In einem weiteren Schritt wurden die Problemfelder im Prozess der Evaluierung von Ideen näher betrachtet. Während bei My Starbucks Idea, Dell Ideastorm und LEGO Ideas eine hohe Anzahl an generierten Beiträgen auffallend ist, weist SAPiens und Tchibo Ideas bisher noch keine bzw. eine beschränkte Themenkategorisierung der Beiträge auf.

Im anknüpfenden Prozessschritt, *Suggestion*, wurde der abgeleitete Prozess der Ideenevaluation und die identifizierten Probleme im Rahmen von Interviews mit

Crowdsourcing Experten diskutiert. Die identifizierten Probleme beziehen sich hauptsächlich auf die Analyse Phase des Ideenevaluierungsprozesses.

In den mit den Experten geführten Interviews wurden die zuvor identifizierten Probleme bestätigt und konkretisiert. Die Experten kamen dabei zu dem Schluss, dass vor allem die Vielzahl an Beiträgen und fehlende Themenkategorisierung die Unternehmen vor große Probleme stellt. Durch die subjektive Betrachtung einerseits und das arbeits- sowie zeitintensive manuelle Vorgehen andererseits werden Organisationen an Kapazitätsgrenzen geführt. Der Einsatz von automatisierten Auswertungsverfahren im Sinne von Text Mining Tools war den Experten bei Ideencommunities nicht bekannt. Ferner konnte in den Experteninterviews erarbeitet werden, welche Phasen im bisherigen Ideenevaluationsprozess das höchste Potenzial für die Unterstützung durch ein TextMining Tool bergen.

Abbildung 2. Unterstützende Funktion eines Text Mining Tool in der Ideenevaluation



Aus den Interviews mit den Experten lässt sich dabei ableiten, dass ein Text Mining Tool insbesondere in der Analysephase einen strukturierten bzw. gesamtheitlichen Überblick über die Ideen darlegt. Indirekt unterstützt das Tool die Entscheidungsphase, da Unternehmen zusätzliche Informationen erhalten, die Unternehmen bei der Wahl der geeigneten Ideen helfen.

Nach der Identifizierung potenzieller Einsatzgebiete und Unterstützungspotenzials eines Text Mining Tools im Rahmen des Prozesses der Ideenevaluierung wurde in der *Development*-Phase in Zusammenarbeit mit Text Mining-Experten die konkrete Ausgestaltung eines Text Mining Tools, in Form von Design-Guidelines, erarbeitet.

Tabelle 1. Darstellung konkreter Probleme und dazugehörige Design-Guidelines

Kategorien	**Probleme**	**Design-Guidelines**
Dokumenten-aufbereitung	Unaufbereitete textuelle Daten	Tool soll Textdokument bereinigen.*
Themen-kategorisierung	Vielfalt an aufbereiteten textuellen Daten	Tool soll Schlüsselwörter extrahieren (Keyword Extraction), deren Häufigkeiten ermitteln und anschließend gewichten.

	Überflüssige/identische Informationen	Tool soll identische Ideen eliminieren.
	Unterschiedliche Formulierungen einer identischen Idee	Tool soll identische Themen-bereiche erkennen und entsprechende Oberkategorien bilden.
Priorisierung	Vielfalt an gegensätzlichen User-Bewertungen	Tool soll Bewertungstendenzen innerhalb der Themenbereiche erkennen (Sentiment-Analyse).
	Relevante Informationen erkennen	Tool soll Themenbereiche sowie Ideen innerhalb dieser Bereiche priorisieren/ gewichten.

*Morphologische, syntaktische und semantische Analyse: a) Entfernen von Zahlen, Satz- und Leerzeichen, b) Behandeln von Groß- und Kleinschreibung, Rechtschreibung, c) Entfernen von Stoppwörtern, d) Stemming, e) Erkennen von Satzgrenzen, f) Durchführung des Part-of-Speech-Tagging, g) Bildung von Wortgruppen, h) Durchführung des Parsing und Junking, i) Erkennen von Synonymen, Mehrdeutigkeit der Sprache (Dialekt).

Hierzu wurden in einem ersten Schritt Probleme textueller Daten identifiziert. Die Anforderung der Aufbereitung dieser Daten liegt darin, das Textdokument mit Hilfe einer morphologischen, syntaktischen und semantischen Analyse zu bereinigen. Die Probleme zwei bis vier können der Themenkategorisierung zugeordnet werden. Zunächst liegt der Fokus auf der Vielfalt an textuellen Daten. Hierbei soll das Tool Schlüsselwörter extrahieren, indem durch eine Häufigkeitszählung bestimmte Wörter identifiziert, geclustert und entsprechend gewichtet werden. Des Weiteren sind die unterschiedlichen Formulierungen einer identischen Idee sowie überflüssige als auch identische Informationen in einem Evaluationsprozess als problematisch einzuordnen. Das Tool soll zum einen identische Ideen eliminieren können. Zum anderen soll es identische Themenbereiche erkennen und Oberkategorien bilden. Die Probleme fünf und sechs, der Kategorie Priorisierung zuzuordnen, zielen auf die Vielfalt an widersprüchlichen User-Bewertungen sowie auf das Erkennen von relevanten Informationen ab. Demnach soll das Tool Bewertungstendenzen innerhalb der Themenbereiche erkennen (Sentiment-Analyse) und anschließend die Themenbereiche sowie Ideen innerhalb dieser Bereiche priorisieren bzw. gewichten können.

Anhand der angesetzten Strukturierung der Probleme konnte ein Prozess mit Anforderungen aufgestellt werden. Diesen gilt es zu erfüllen, um das gesetzte Ziel einer automatisierten und effizienten Ideenevaluation zu erreichen.

4 Geplante weitere Schritte für die Entwicklung eines Text Mining Tools

In dem vorherigen Abschnitt haben wir die vorläufigen Ergebnisse in Bezug auf die notwendigen Gestaltungsanforderungen eines Text Mining Artefakts konkret herausgearbeitet. Die geplanten weiteren Schritte für die Entwicklung dieses Text Mining Tools in Anlehnung an den DSR-Ansatz sehen wie folgt aus: In der vierten Phase, *Evaluation*, werden die Gestaltungsanforderungen durch zusätzliche Interviews mit Plattformbetreibern bezüglich der potenziellen Anwendung bzw. Nutzbarkeit analysiert. Ferner wird das Tool entwickelt und in einem abschließenden Schritt erneut evaluiert. Das gewählte Vorgehen zielt darauf ab, den Prozess der Ideenevaluation effizienter zu gestalten und folglich die Nutzung bzw. Verwertung der Ideen aus Ideencommunities zu unterstützen. Aus diesem Grund wird das Tool auf Effizienz im Sinne von Geschwindigkeit untersucht. Mit Hilfe einer einfachen Zeitmessung könnte man die benötigte Dauer des Ideenevaluationsprozesses durch Anwendung des Tools ermitteln und anschließend dem zweistufigen manuellen Prozess durch User und Experten gegenüberstellen.

5 Zusammenfassung

In dem vorliegenden Paper haben wir uns mit der Gestaltung eines Artefakts auf Grundlage des DSR-Ansatzes nach Vaishnavi und Küchler [13] auseinander gesetzt. Das Ziel hierbei war es die Phasen der Ideenevaluierung näher zu beleuchten, Unterstützungspotenziale durch ein Text Mining Tool in den einzelnen Phasen zu identifizieren und auf Basis dessen die konkreten Anforderungen an ein solches Tool abzuleiten. Der sich anschließend hieraus ergebene Beitrag bezieht sich auf die zwei Bereiche Praxis und Wissenschaft.

Zum einen liegt der praktische Mehrwert unseres Beitrags darin Unternehmen Handlungsempfehlungen zum effizienten Einsatz und zur Implementierung eines Text Mining Tools zur Unterstützung der Ideenanalyse in Ideencommunities zu geben. Die Vorteile einer solchen Lösung liegen auf der Hand. So erlaubt die Implementierung einer Text Mining Lösung dem Plattformbetreiber die Ausschöpfung von Automatisierungspotenzial, das zu internen Ressourceneinsparungen in Form von Kosten und Arbeitskapazität führt. Zusätzlich bietet eine Text Mining Lösung den Unternehmen die Erkennung von neuen, bisher nicht entdeckten sowie relevanten Zusammenhängen und dient diesen folglich als Informations- und Entscheidungsunterstützung. Folglich kann das vorhandene Potenzial der vorliegenden Daten besser ausgeschöpft werden.

Der theoretische Mehrwert unseres Beitrags liegt in der Erweiterung des Erkenntnissstandes der bestehenden Crowdsourcing Literatur. Dieser wird einerseits erreicht durch die gezielte sowie der tiefergehende Betrachtung und Analyse der Ideenevaluierung im Crowdsourcing. Bisherige Studien zu dem Thema beschäftigten sich primär damit wie sich die Crowd in die Evaluation von Ideen einbinden lässt [38] und welche Individuen sich am besten für die Bewertung von Ideen eignen [39]. Zum anderen wird die bestehende Literatur durch die in dieser Studie vorgestellten Gestaltungsempfehlungen für ein Text Mining Tool um eine neue Forschungsperspektive erweitert. Unser Beitrag liefert dabei eine Design Theorie [40] zur Gestaltung eines Textmining Artefakts zur effizienten Unterstützung der Ideenevaluation im Crowdsourcing.

Literaturverzeichnis

1. Zogaj, S., Bretschneider, U., Leimeister, J.M.: Managing Crowdsourced Software Testing – A Case Study Based Insight on the Challenges of a Crowdsourcing Intermediary. Journal of Business Economics (JBE) (DOI: 10.1007/s11573-014-0721-9) (2014)
2. Poetz, M.K., Schreier, M.: The value of crowdsourcing: can users really compete with professionals in generating new product ideas?. Journal of Product Innovation Management, 29(2), 245-256 (2012)
3. Hienerth, C., Lettl, C., Keinz, P.: Synergies among producer firms, lead users, and user communities: The case of the LEGO producer–user ecosystem. Journal of Product Innovation Management, 31(4), 848-866 (2014)
4. Blohm, I., Bretschneider, U., Leimeister, J.M., Krcmar, H.: Does collaboration among participants lead to better ideas in IT-based idea competitions? An empirical investigation. International Journal of Networking and Virtual Organisations, 9(2), 106-122 (2011)
5. Blohm, I., Leimeister J.M., Krcmar, H.: Crowdsourcing: How to Benefit from (Too) Many Great Ideas. MIS Quarterly Executive, 4(12), 199-211 (2013)
6. Majchrzak, A., Malhotra, A.: Towards an information systems perspective and research agenda on crowdsourcing for innovation. The Journal of Strategic Information Systems, 22(4), 257-268 (2013)
7. Blohm, I., Leimeister, J.M., Rieger, M., Krcmar, H.: Controlling von Ideencommunities – Entwicklung und Test einer Ideencommunity-Scorecard. Controlling, 2(23), 96-103 (2011)
8. Cohen, W.M., Levinthal, D.A.: Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly, 35(1), 128-152 (1990)
9. Roberts, N., Galluch, P., Dinger, M., Grover, V.: Absorptive Capacity and Information Systems Research: Review, Synthesis, and Directions for Future Research. MIS Quarterly, 36(2), 625-648 (2012)
10. Malone, T.W.: Modeling coordination in organizations and markets. Management science, 33(10), 1317-1332 (1987)
11. Hippner, H., Rentzmann, R.: Text mining. Informatik-Spektrum, 29(4), 287-290 (2006)
12. Yang, J., Adamic, L.A., Ackerman, M.S.: Crowdsourcing and knowledge Sharing: Strategic user behavior on taskcn. Proceedings of ACM Electronic Commerce'08, 246-255 (2008)
13. Vaishnavi, V., Kuechler, W.: Design research in information systems. (2004)
14. Schreyögg, G., Schmidt, S.: Absorptive Capacity – Schlüsselpraktiken für die Innovationsfähigkeit von Unternehmen. WiSt Heft 10, 474-479 (2010)
15. Lichtenthaler, U., Ernst, H.: Opening up the innovation process: The Role of Technology Aggressiveness. R&D Management, 39(1), 38-54 (2009)
16. Zahra, S.A., George, G.: Absorptive capacity: A review, reconceptualization, and extension. Academy of Management Review, 27(2), 185-203 (2002)
17. Todorova, G., Durisin, B.: Absorptive capacity: Valuing a reconceptualization. Academy of Management Review, 32(3), 774-786 (2007)

18. Lane, P.J., Koka, B.R., Pathak, S.: The Reification of Absorptive Capacity: A Critical Review and Rejuvenation of The Construct. Academy of Management Review, 31(4), 833-863 (2006)
19. Ebner, W., Leimeister, J.M., Krcmar, H.: Community Engineering for Innovations: The Ideas Competition as a method to nurture a Virtual Community for Innovations. R&D Management, 39(4), 342-356 (2009)
20. Leimeister, J.M.: Einführung in die Wirtschaftsinformatik. Springer-Verlag. (2015)
21. Hearst, M.: What is text mining. SIMS, UC Berkeley (2003)
22. Feldman, R., Dagan, I.: Knowledge Discovery in Textual Databases (KDT). KDD., 95, 112-117 (1995)
23. Hippner, H., Rentzmann, R.: Text mining. Informatik-Spektrum, 29(4), 287-290 (2006)
24. Sullivan, D.: Document warehousing and text mining: techniques for improving business operations, marketing, and sales. John Wiley & Sons, Inc. (2001)
25. Hotho, A., Nürnberger, A., Paaß, G.: A brief survey of text mining. Ldv Forum, 20(1), 19-62 (2005)
26. Carstensen, K.U., Ebert, C., Ebert, C., Jekat, S., Langer, H., Klabunde, R.: Computerlinguistik und Sprachtechnologie: Eine Einführung. Springer-Verlag (2009)
27. Feldman, R., Sanger, J.: The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge University Press (2007)
28. Williams, S.: A survey of natural language processing techniques for text data mining (2000)
29. Miner, G.: Practical text mining and statistical analysis for non-structured text data applications. Academic Press (2012)
30. Aggarwal, C.C., Zhai, C.: A survey of text classification algorithms. Mining text data. Springer US, 163-222 (2012)
31. Weiss, S.M., Indurkhya, N., Zhang, T.: Fundamentals of predictive text mining, 41. London: Springer (2010)
32. Konchady, M.: Text mining application programming. Charles River Media, Inc. (2006)
33. Tan, S.: Neighbor-weighted k-nearest neighbor for unbalanced text corpus. Expert Systems with Applications, 28(4), 667-671 (2005)
34. Hotho, A., Nürnberger, A., Paaß, G.: A brief survey of text mining. Ldv Forum, 20(1), 19-62 (2005)
35. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. Journal of management information systems, 24(3), 45-77 (2007)
36. Simon, H.A.: The sciences of the artificial. MIT press. (1996)
37. March, S.T., Smith, G.F.: Design and natural science research on information technology. Decision support systems, 15(4), 251-266 (1995)
38. Leimeister, J.M., Huber, M., Bretschneider, U., & Krcmar, H.: Leveraging crowdsourcing: activation-supporting components for IT-based ideas competition. Journal of management information systems, 26(1), 197-224 (2009)
39. Poetz, M.K., & Schreier, M.: The value of crowdsourcing: can users really compete with professionals in generating new product ideas? Journal of Product Innovation Management, 29(2), 245-256 (2012)
40. Gregor, S., & Jones, D.: The anatomy of a design theory. Journal of the Association for Information Systems, 8(5), 312 (2007)
41. Kruse, P., Schieber, A., Hilbert, A., & Schoop, E.: Idea Mining–Text Mining Supported Knowledge Management for Innovation Purposes (2013)

Assessing the Potential Value of Software-Support for the Venture Creation Process

Pascal Stegmann[1], Marin Zec[1] and Florian Matthes[1]

[1] Technische Universität München, Chair for Informatics 19 (sebis), Munich, Germany
{pascal.stegmann,marin.zec,matthes}@tum.de

Abstract. Many newly founded companies fail due to a lack of experience of the founders, a missing connection to mentors, investors or other valuable contacts, or building a product without a real market need. Formalized frameworks like Lean Startup suggest that there is a structured way of evaluating ideas and building companies that increase the chances of success. In this paper we investigate and identify potential for software support in managing new venture creation. We conducted 10 semi-structured interviews with employees at business incubators and startup founders. Based on our observations we derived high-level guidelines that need to be taken into account for a software to provide added value for management activities of venture creation processes.

Keywords: Entrepreneurship, Venture Creation Process, Knowledge Management, Social Software

1 Introduction

Evaluating an idea for a business model and ultimately building a company is a complex and high risk undertaking. After earlier research on the different stages of venture creation [1] which suggested a predictable, chronological venture creation process, more recent research acknowledges the complex, dynamic and often unpredictable nature of the venture creation process [2], [3]. Focusing on the early phase, i.e. the process between initial conception of an idea until a product or service is sold to customers, Bhave [4] developed a process model with distinct sequential steps while at the same time emphasizing the iterative and non-linear character of the process. Abstract models like this help to isolate and understand different factors increasing the chances for success like having experience [5], operating in a favorable environment [6] or having a strong social network and personal relationships [7-12] but are not necessarily helpful for founders as they do not provide a prescriptive model.

By combining multiple current management trends such as Agile Development, Lean Thinking and Customer Development, various methods emerged in the last years aiming to provide guidance and a conceptual framework to manage the complexity of the venture creation process.

For instance, the Lean Startup method [13] promotes an iterative process involving short feedback loops between founder(s) and potential customers. The underlying

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Stegmann, P.; Zec, M.; Matthes, F. (2017): Assessing the Potential Value of Software-Support for the Venture Creation Process, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1417-1428

assumption of the Lean Startup method and similar approaches is that there is some sort of a prescriptive blueprint process model or best-practice patterns such as certain practices or artifacts for creating a new business. Given this formalization of the venture creation process, the topic of software support becomes more and more relevant to support founders in implementing these frameworks, increase transparency in the process and support communication amongst involved parties.

The objective of this paper therefore is to investigate and identify potential for software support in managing new venture creation, thereby providing a basis for future research on software support of the venture creation process.

To prevent potential premature generalizations due to conceptual bias towards a particular potentially narrow conceptual framework and existing preconceptions about venture creation processes, we chose an inductive, qualitative research approach.

2 Research Approach

We conducted 10 semi-structured interviews with startup founders and employees at different business incubators. The interviews centered around the following four themes:

1. Which general characteristics describe the venture creation process?
2. How is knowledge transferred and a lack of experience compensated in this context?
3. How are tools and artifacts used over the course of the venture creation process?
4. What characterizes the social exchange during the venture creation process (e.g. networking and access to experts)?

2.1 Research Design & Method

We followed an exploratory approach based on 10 semi-structured interviews. This way, we were able to cover common questions and themes as well as remain open to issues brought forward by our interview partners. The interviews were conducted in two rounds with two types of interview partners (see section 2.2 for more details).

As suggested by Mayring [14], the interviews were transcribed as detailed as necessary to capture important information and enable further processing. The transcriptions were coded using the four themes mentioned above as codes (i.e. *general characteristics*, *knowledge exchange*, *tools/artifacts* and *social exchange*). To end up with the final observations we followed an approach presented by Miles and Huberman [15], by marking certain patterns in the extracted text segments and subsume particular aspects into more general observations.

In a final step, we abstracted from the descriptive observations to derive prescriptive guidelines for venture creation management software which address one or more of the observations. This was done by collecting underlying problems for each observation. By grouping these problems, we could again subsume particular aspects of these problems into general problems that a software support needs to address. These general aspects consequently represent the design guidelines.

2.2 Sample Selection & Description

The interviews were conducted in two rounds using different interview guidelines to account for the different characteristics of the types of interview partners (see table 1). The first round involved employees at business incubators since we assumed that business incubators tend to follow a (semi-)structured approach to company building and supporting the venture creation process of startups. Further, we assumed that employees at business incubators would be likely to observe and recognize potential patterns of startup founder experiences since they interact with multiple startups.

Table 1. Description of Interview Partners

	No	*Code*	*Role*	*Description*
Round 1	1	PM1	Project Manager	Publicly funded incubator program, thematic focus, 1-5 years old
	2	IL	Innovation Lead	Corporate incubator/research lab, industry focus, 1-5 years old
	3	PM2	Program Manager	Publicly funded incubator program, industry focus, <1 year old
	4	COO 1	Chief Operating Officer	University affiliated accelerator program, technology focus, <1 year old
	5	CM	Community Manager	University affiliated incubator program, focus on scalable business models, >5 years old
	6	COO 2	Chief Operating Officer	Independent company builder, focus on scalable business models, >5 years old
	7	CEO	Chief Executive Officer	Independent accelerator program, focus on scalable business models, <1 year old
Round 2	8	TF	Technical Founder	Platform in e-commerce space, before scaling phase, received first financing round
	9	BF1	Business Founder	Media startup, validation phase with internal beta test, no funding
	10	BF2	Business Founder	Media startup, validation phase with public beta test, no funding

To pick the interview partners of business incubators, a list of business incubators in Germany containing 58 entries was created. Subsequently, the initial list was reduced to a shortlist of 22 business incubators that covers a broad spectrum of possible approaches to business incubation. Business incubators on the shortlist were contacted, from which seven ultimately agreed on participating in the research project. The business incubation programs differed in terms of age of the program, industry focus, objective of the organization, phases of admitted startups as well as duration and degree of involvement of the business incubator.

In a second round we interviewed startup founders to get reports on first-hand experience with entrepreneurship. We focused on first time founders to get insights into the problems of unexperienced founders as they would benefit most from software support. In terms of business models, we made sure that there is a significant amount of technology involved and frameworks such as Lean Startup are potentially applicable.

3 Observations

We found several challenges and characteristics that define the environment of a potential system support. In this section, we present our observations structured according to the four themes of interview questions, presented in our research approach (*general characteristics*, *knowledge exchange*, *tools/artifacts* and *social exchange*)[1].

3.1 General Characteristics

An important observation influencing the venture creation process is the resource restrictions startups face. Their critical resource is most often available time. Thus, many activities fall short because of lower priority and due to low perceived value for the startup. Founders have to take a pragmatic approach and as mentioned in the interviews, unnecessary tasks or too much formalization and fixed processes may produce too much overhead that is not valuable and are therefore rejected.

O1 - ***Founders face resource restrictions and have only limited time available.*** (2 of 10; COO2, PM2)

It was often mentioned that each startup is unique and as one interview partner put it *"there is no blueprint"* (COO2) of how to build a company. Given the fact that by definition a startup is doing something new and innovative and further considering the different influencing factors, the result is often an unstructured, chaotic and fast changing environment with low predictability of outcomes. There is seldom a protocol or process to follow, but rather a pragmatic approach has to be taken to handle the high degree of uncertainty a startup faces. This manifests itself in a required high degree of flexibility, how a process evolves, which tasks to perform, which information to capture, etc.

O2 - ***Each startup is unique and requires a high degree of flexibility with regards to the performed tasks and processed information.*** (6 of 10; CM, COO1, COO2, IL, PM2, TF)

As mentioned by employees at business incubators, providing support for the founders is necessary and helpful, but the initiative ultimately has to come from the founders, i.e. an advisor cannot push a team to be successful if they do not have the drive to achieve success themselves. Most often the support is therefore on a request basis, as one

[1] The number in parentheses represents the number of interviews this observation is based upon, i.e. 7 of 10 means that this was mentioned by 7 interview partners out of the total 10 interviewed persons. The codes in parentheses refers to the interview codes as introduced in the methodology section in table 1 and indicate the source of information.

incubator mentioned "*we see ourselves as a service provider for the startups and only give recommendations, the decision has to come from the startup itself*" (COO1). This also seems to be an important trait to be successful in the long-term and the alternative can be counterproductive if a team is just executing what someone else tells them. Founders themselves mentioned they want to remain in control and drive the development and not be forced by a process or third party to do something or use some tools they do not want. This has to be balanced as they still have to be receptive to feedback and ultimately make the right assessment.

O3 - *Founders want and need to take the initiative and have the deciding power.* (7 of 10; CM, COO1, COO2, IL, PM1, PM2, TF)

3.2 Knowledge-Related Observations

Given the complexity of the venture creation process and the required tasks at hand it is hardly possible to acquire all the required knowledge in advance. Especially unexperienced founders lack process knowledge, i.e. what to do and how to do or approach things. As mentioned, they try to use their own judgment what to do or try to educate themselves, but often due to a lack of time (see O1) it is not possible to perform an elaborate research on the best practice for a specific task. They rather resort to trial and error and incorporate gained experience and knowledge over time. Even experienced founders face situations where they may not be knowledgeable due to the uncertainty of the process.

O4 – *Founders often lack relevant knowledge about the venture creation process and best practices.* (4 of 10; BF1, CEO, COO1, IL)

The most valuable knowledge is often tacit and, thus, hard to explain or codify. Such kind of knowledge is built from experience and for example comprised of the ability to assess a situation with limited information and still provide a promising recommendation. It is often not possible to break this decision making process down into simple and generally applicable rules, due to the complexity of influencing factors that have to be taken into account. This is a value add of advisors and mentors and why the personal interaction and support by experienced people is considered to be important in this context. One founder emphasized that they did not lack any support in the beginning besides advisors: "*we could have needed that (advisors) from the beginning ... someone who raps our knuckles ... we underestimated this*" (TF). Experienced-based knowledge might be a differentiating factor as it is difficult to replicate.

O5 - *Valuable knowledge is often tacit and based on experience.* (4 of 10; BF1, COO1, COO2, TF)

Besides tacit knowledge, certain information on specific topics is documented and accessible via the internet. The challenge is often to make an effort to find it. As mentioned, the scattered information makes it burdensome to find a specific information and assess the validity and applicability in a given situation.

O6 - *Certain information and knowledge is spread across different sources and difficult to find.* (3 of 10; CEO, CM, TF)

Many challenges are unique to each startup and depend on the specific context (e.g. product, business model or degree of innovation). However, it was mentioned that overall there are also problems that are recurring across startups. A startup may only face this problem once or a few times and therefore sees no need to document such knowledge. Although it could be helpful for other startups there is no real incentive to share this. An example could be how to set up a legal structure, which is only done once in the beginning. Such knowledge is often provided by people interacting with multiple companies like mentors or incubators that can transfer such knowledge between companies. As an example, incubators use strategy days or simple lists describing what to consider when founding a company to distribute this knowledge.

O7 - *Certain problems are recurring between startups but not necessarily within a single startup.* (4 of 10; CM, COO2, PM1, PM2)

Different types of knowledge exist and are subject to different frequencies of change. Legal topics were given as an example for quite static knowledge. As a contrast, topics like online marketing or new emerging technologies were mentioned as topics that are subject to frequent changes, which makes it difficult for documented knowledge to remain relevant. The difficulty therefore becomes to balance the efforts of collecting and formalizing such knowledge with the created value of doing so.

O8 - *For certain domains, new knowledge is created often or existing knowledge is subject to frequent changes.* (3 of 10; CM, COO1, COO2)

3.3 Tools- and Artifacts-Related Observations

Founders tend to feel overwhelmed by the amount of offered software tools for the various tasks they need to perform like prototyping, business modelling, analytics, etc. They have difficulties identifying relevant and useful ones. One founder mentioned: "*I think there are too many tools, we are testing tools all the time*" (BF2). Some software tools are trending or considered standard in certain areas. However, it remains up to the founder to assess the suitability for a given situation.

O9 – Multiplicity of software tools for the various tasks of a founder are available, which requires a continuous assessment of suitability. (5 of 10; BF2, CM, COO1, PM2, TF)

Various artifacts are created and used over the course of the venture creation process. As an example, for business modeling purposes a pitch deck, Business Model Canvas or some form of business plan is created but mostly by using generic data processing software that does not provide much structure. Therefore, the structuring needs to be provided by the user, e.g., by using other example documents as templates.

O10 - ***Multiple types of artifacts are created over the course of the venture creation process. However, little structural support is provided as mostly generic software is used.*** (5 of 10; BF1, BF2, CEO, COO1, TF)

Some artifacts such as the Business Model Canvas or a business plan are perceived as a standard amongst startups. For some business incubators they are mandatory to assess the business model: "*In the beginning we have an assessment center, where we use a Business Model Canvas*" (COO1). However, sometimes there seems to be a lack of experience of how to create and work with these artifacts. One founder mentioned, he used the Business Model Canvas but "*besides filling things in boxes … it did not provide much value*" (BF1). The idea of most artifacts is to convey a certain thinking about a problem by providing a structure. However, if people do not know how to effectively use this structure or are not provided with any form of guidance the artifacts lose their purpose.

O11 - ***Founders do not necessarily know how to effectively use certain artifacts.*** (3 of 10; BF1, IL, TF)

Pen-and-paper tools seem to play an important role for specific artifacts and project management purposes. Value is still perceived in using these forms instead of working digitally only. Especially in agile methods and when applying the Lean Startup approach, the use of posters of boards and special artifacts is prevalent as supported by the empirical data. It is more suitable for group work and in-person discussions, due to better visibility. However, there remains a gap when trying to integrate the physical pen and paper artifact into a digital workflow. As mentioned by one founder, the pen and paper versions get outdated ("*the one (Business Model Canvas) as it is hanging on the wall is not up-to-date anymore*", BF2) and lose its usefulness if the form is switched from physical to digital.

O12 - ***Founders see value in pen-and-paper tools, but it is difficult to keep physical and digital versions in sync.*** (3 of 10; BF2, CEO, TF)

3.4 Social Exchange-Related Observations

Due to the knowledge often being tacit and advice being contextual, various types of interaction for knowledge transfer are required, such as personal meetings and discussions. As was mentioned in the interviews, the exchange benefits from a more informal setting and that is often why incubators organize events or try to have startups be co-located to foster exchange, as it increases the willingness to share information. As one incubator mentioned: "*The teams being co-located proved to be a major* *strength, as they are communicating a lot amongst each other … we were almost surprised how well it works*" (PM2).

O13 - *Exchange of knowledge and experience is often informal, i.e. through relatively unstructured personal interaction between peers and advisors/mentors.* (8 of 10; BF1, CM, COO1, COO2, IL, PM1, PM2, TF)

Most valuable knowledge is often tacit and held by experts, such as experienced founders or investors. Due to their expertise, these experts are usually sought out by many people seeking for help and are therefore difficult to get in touch with. Even before an initial contact, it may be difficult to identify the right expert who might be a good fit, as there is not always transparency about the area of expertise and actual knowledge that the person could provide. As mentioned by business incubators, this matchmaking process is one of the most important value adds, i.e. identifying the right fit and giving introductions to overcome this barriers founders usually face. They usually provide this through alumni networks or other forms of partner networks.

O14 - *Founders often lack access to experienced people like mentors/experts/etc.* (7 of 10; CM, COO1, COO2, IL, PM1, PM2, TF)

In interaction with advisors, mentors and other people that need deep insights into the company to give useful advice, sensitive information is shared. Especially in the beginning when there is often not more than an idea, founders fear revealing proprietary information and get their intellectual property stolen. As one interview partner mentioned: "*startups need to have trust in their mentors … startups have to decide themselves who they can trust and who they want to share it (information) with*" (CM). Founders may be hesitant to document and share certain information and knowledge, as they want to have control over who has access to it. In a similar way, advisors might share confidential material too, to provide examples of documents or information that might be helpful for founders but need to be kept confidential.

O15 - *Sensitive information needs to be handled and bears the fear of founders of revealing proprietary information.* (3 of 10; CM, COO1, TF)

4 Design Guidelines

To address the mentioned observations, design guidelines for the IT support of the venture creation process were developed and are presented in this section. The reference code in parentheses indicates the observations the guideline is based upon.

DG1 - *Integrate system usage into existing workflows:* Due to the high resources restrictions startups face, additional work needs to be minimized in order to increase the acceptance and adoption of a tool. The effort to use a system can be minimized by integrating system usage into existing workflows and not add additional process steps. (O1)

DG2 - ***Enable contextual accessibility of knowledge:*** On a similar note, to minimize the barrier and required effort of switching between learning and doing something, relevant knowledge should be provided in a specific context, i.e. where and when it is needed or should be applied. The alternative would be to provide all available knowledge in a central location thereby having increased switching costs to access the knowledge. (O1)

DG3 - ***Provide functionality to set goals:*** Given the complexity of the venture creation process, focusing too much on individual steps can be counterproductive. Instead, by supporting the setting of goals and milestones collaboratively with stakeholders, accountability can be created without compromising on the flexibility of the actual implementation compared to a fixed defined process. (O2, O3)

DG4 - ***Enable structuring flexibility:*** To further account for the complexity of the process and not be too restrictive, flexibility and adaptability of the system needs to be ensured, with regards to captured data, information and structuring capabilities of the process to address the individuality of the startup. (O2, O3, O10)

DG5 - ***Provide support on a suggestion basis:*** Any form of decision support needs to remain on a suggestion basis. Control and final implementation needs to remain with founders to increase the acceptance of a support by not restricting the founder's freedom of action. As an example, possible next steps in a certain situation could be suggested based on collected data but which actions to implement ultimately has to be decided by a human. (O2, O3)

DG6 – ***Support bridging medium gap between analog and digital tools:*** Given the perceived value of analog tools and artifacts, bridging the medium gap between analog and digital tools should be enabled to combine the benefits of both forms, i.e. ease of creation and interaction of analog forms with ability to share and collaborate across locations of digital forms. As an example, results of a brainstorming on a flipchart could be captured with a camera and further processed in a digital workflow. (O12)

DG7 – ***Enable knowledge emergence through usage:*** Given the resource restrictions, taking time to formalize created knowledge is not necessarily the highest priority. Therefore, the emergence of knowledge and best practices through the use of the system should be enabled to minimize the effort of knowledge explication. This could be achieved through providing structuring flexibility (see DG4) and afterwards analyze the usage for patterns to use as suggestions for the future. (O1, O5, O8)

DG8 – ***Incentivize knowledge explication:*** In addition to automate some form of knowledge creation, explication and sharing of knowledge needs to be incentivized to keep the knowledge base up-to-date with valid knowledge and thereby relevant for the founder. (O1, O5, O8)

DG9 – *Provide shared knowledge base with common information:* A shared knowledge base with relevant information for the venture creation process (e.g., best practices for common processes, suggested tools for use cases, etc.) to compensate for the difference in knowledge and establish a common understanding needs to be provided. As mentioned, this could be a central repository or wiki to collect all knowledge, though providing context for the knowledge is important (see DG2). (O4, O6, O7, O9, O11)

DG10 – *Ensure adaptability of knowledge:* As some knowledge might be changing it needs to be adapted to stay relevant and provide value to the users. Mechanisms need to be provided to easily adapt the knowledge base to account for the changing nature of information and knowledge while still ensuring the validity of the knowledge, possible through some form of version control and consensus mechanism. (O8)

DG11 – *Provide functionality to share information amongst stakeholders:* To foster collaboration amongst stakeholders easy sharing of information with other stakeholders should be enabled to lower the barrier of knowledge transfer and simplify creating a context for discussion. This could be achieved through providing access rights to certain information collected in a system or using of existing communication channels to distribute information. (O1, O13)

DG12 – *Foster social exchange and relationship building:* Given the mentioned barrier to formalize and capture certain knowledge, an additional way of knowledge exchange is through social interaction. Therefore, social exchange and engagement between users needs to be incentivized to support relationship building and improve the knowledge transfer of tacit knowledge. This could be achieved by encouraging the initial interaction amongst users and further through lowering the technical barriers of engagement, i.e. make it as easy as possible to connect and interact with each other. (O5, O13)

DG13 – *Simplify identification of experts with relevant knowledge:* As mentioned, there is often a lack of access and possibly a lack of transparency about the capabilities of a certain person. Therefore, the identification and access to people with relevant knowledge and expertise should be simplified to enhance the matching process and reduce the necessity for human intervention. (O14)

DG14 - *Create a digital space providing trust and confidentiality:* As confidential information is often handled in this context, formalizing and submitting such information to a third party system could pose a high barrier to system usage. To overcome this issue a space of trust and confidentiality needs to be created by giving transparent access control to the data owner to support the willingness of users to share information. (O15)

5 Conclusion

At a rather abstract level, common features between different venture creation processes could be identified such as the use of a business plan or Business Model Canvas. However, there are considerable differences in the way of usage and perceived value of applying certain methods or using certain types of artifacts. From the perspective of our interview partners, venture creation processes are perceived to be idiosyncratic. Furthermore, their application and usage does not necessarily seem central to success.

Our study suggests that the venture creation process is dominated by three characteristics: (1) the unpredictability in the way it evolves over time, (2) a (perceived) lack of knowledge and experience among most founders and (3) its social nature to get access to valuable resources and enable transfer of experienced-based knowledge between different stakeholders. According to our interview partners, knowledge and social exchange may compensate for a lack of formalization and are perceived to be increasing the likelihood of success.

In the present study, we derived guidelines based on our observations to inform the design of software systems for venture creation management. While the focus of this study was to gain insights into entrepreneurial practice, we acknowledge the yet hypothetical nature of the guidelines. Given the limited sample, a thorough validation and possible adjustment of the derived design guidelines is subject to future research.

Looking at different paradigms as a basis for such software support, Business Process Management (BPM) comes to mind as a way to support the management of the venture creation process. However, given our empirical findings, the high uncertainty and contextual dependency of the process, the focus on predefined process steps does not seem like a viable solution.

The uncertainty of the process and the context dependency on the knowledge and decision of the user calls for a more social and flexible approach such as Adaptive Case Management (ACM) seems more appropriate. It aligns with the goals driven approach presented in the guidelines and the focus on artifacts that are used in the process. It further emphasizes collaboration and social interaction rather than a step-by-step process. Still further research should provide a thorough discussion and investigation of the suitability of this paradigm.

Given the multitude of involved stakeholders, we think it is important to focus on those who would benefit and/or contribute the most to reduce the complexity. However, there seems to be an asymmetry of incentives between those benefitting and those contributing to such a system. We see two main directions to extend the research: 1) identification of stakeholders who would benefit most from software support and 2) research on incentive mechanisms which would drive engagement towards and adoption of a venture creation management system.

References

1. Ruhnka, J. C. and Young, J. E.: A Venture Capital Model of the Development Process for New Ventures. Journal of Business Venturing 2(2), 167-184 (1987)
2. Liao, J., Welsch, H., and Tan, W.-L.: Venture gestation paths of nascent entrepreneurs: Exploring the temporal patterns, Journal of High Technology Management Research 16(1), 1-22 (2005)
3. Wright, M. and Marlow, S.: Entrepreneurial activity in the venture creation and development process. International Small Business Journal 30(2), 107-114 (2012)
4. Bhave, M. P.: A process model of entrepreneurial venture creation. Journal of Business Venturing 9(3), 223-242 (1994)
5. Delmar, F. and Shane, S.: Does Experience Matter? The Effect of Founding Team Experience on the Survival and Sales of Newly Founded Ventures. Strategic Organization 4(3), 215–247 (2006)
6. Isenberg, D.: What an Entrepreneurship Ecosystem Actually Is. Harvard Business Review. , https://hbr.org/2014/05/what-an-entrepreneurial-ecosystem-actually-is (Accessed: 06/01/2016)
7. Hayter, C. S.: Conceptualizing knowledge-based entrepreneurship networks: perspectives from the literature. Small Business Economics 41(4), 899–911 (2013)
8. Jiang, X., and Yanqiu, L.: Impacts of Entrepreneurial Network on Entrepreneurial Learning. In: 2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering (pp. 401–404). Kunming, China (2010)
9. Leyden, D. P., Link, A. N., and Siegel, D. S.: A theoretical analysis of the role of social networks in entrepreneurship. Research Policy, 43(7), 1157–1163 (2014)
10. Panda, S., and Dash, S. K.: Friends and strangers: leveraging the power of networks for new venture success. Development and Learning in Organizations: An International Journal 29(2), 6–9 (2015)
11. Stam, W., Arzlanian, S., and Elfring, T.: Social capital of entrepreneurs and small firm performance: A meta-analysis of contextual and methodological moderators. Journal of Business Venturing 29(1), 152–173 (2014)
12. Sullivan, D. M., and Marvel, M. R.: Knowledge Acquisition, Network Reliance, and Early-Stage Technology Venture Outcomes. Journal of Management Studies 48(6), 1169–1193 (2011)
13. Ries, E.: The Lean Startup. Crown Business, New York (2011)
14. Mayring, P.: Einführung in die Qualitative Sozialforschung. Beltz Verlag, Weinheim (2002)
15. Miles, M. B., and Huberman A. M.: Qualitative Data Analysis. Sage Publications, Thousand Oaks (1994)

TOOLS FOR OPEN SERVICE DESIGN
An Analysis into Existing and Future Methodologies

Andrew Cooney[1], Timo Zink[1], Dr. Julia Jonas[1]

[1] Friedrich Alexander Univeristy, Information Systems 1, Erlangen Nuremberg, Germany
{andrew.cooney, timo.zink, julia.jonas}@fau.de

Abstract. The growing importance of the service industry in today's information technology-oriented domain is indubitable. In contrast to the manufacturing industry, which has had time to develop, mature and standardise processes, there exists in the service industry a dearth of standardised design methodologies. In analysing an extent of relevant new literature, this illustrative case study examines three service design tools chosen on the basis of their applicability to open innovation. One tool is thus selected and applied to a currently existing, but hitherto under used, technology. The purpose of this illustrative case study is to expand the service design repertoire of service innovation managers in the field of open innovation.

Keywords: Open Service Design, Prototyping, Open Innovation, Service Innovation, Co-creation

1 Introduction

The importance of the service industry in the global context is a relatively new phenomenon; with a broad definition of services and its components still contested among academic communities. Services comprise more than 70% of aggregate gross domestic product and employment in the Organization for Economic Cooperation and Development countries [24]. Moreover, increased globalisation, in concurrence with the increased access to the internet [5], has accelerated competition and challenged companies to be more agile in light of increasingly complex consumer demands [26].

Open innovation is a concept that is gaining more traction within service design thinking and can help companies achieve the growth in service revenues they seek [10]. A firm achieves this by developing outward-looking strategic approaches to research and development in the attempt to leverage potential value from a broader environment [24]. However, the methodologies in regards to utilising this concept and resulting complexities are, until now, not yet widely investigated within the realm of service design. With the increasing popularity of open innovation, service innovation managers may have to learn how to effectively interact with the consumers for whom they create services.

This paper addresses the challenges and opportunities that service innovation managers may face when approaching open service design. This is achieved through an evaluation of the opportunities and limitations of a set of existing service design tools and recommending the best fitting tool with which service innovation managers can most effectively respond to this challenging environment. Thereafter this recommended design tool is applied to an emerging technology to demonstrate how the challenges and opportunities relating to service design could be minimised and leveraged respectively.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Cooney, A.; Zink, T.; Jonas, J. (2017): TOOLS FOR OPEN SERVICE DESIGN An Analysis into Existing and Future Methodologies, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1429-1440

2 Research Methodology

The work that ensues has made use of an illustrative case study [35] to establish a conceptual discussion of prototyping within open service design. Initially, theory regarding select service design tools is examined in the context of open innovation. Following this the reader is exposed to possible use case studies in which prototyping has been applied. Lastly, future uses of prototyping is posited in the use of currently-available, however not widely used, technology.

3 Theoretical Background of Open Service Design

Until recently limited attention has been paid to developing systems that can address this changing situation, and specifically to the issue of new service development [26]. New internet technologies have not only increased the reach of information but also exposed hitherto non-internet users to a much wider "virtual" market [5]. Consequently consumer demand has evolved to reflect this increasing complexity. As Peter Drucker, in his interview with Chesbrough [10], states "What the customer buys and considers value is never a product. It is always utility – that is, what a product does for him." This trend towards consumer-centricity has pushed companies to engage in new forms of research development; looking at how to integrate external stakeholder knowledge into product and service development [24].

One such approach is open innovation. This recognises and builds on the changes in the dynamics of the consumer's interaction with the company from simply a consumer of pre-determined services to that of value co-creation with the company. By engaging outside knowledge through open innovation, companies can achieve the growth in market share and in revenue that they seek [10]. However, it is often the case that activities performed by managers are inconsistent with the demands of the consumer they try to meet [39]. Consequently, a plethora of tools and methodologies have been developed under the broad umbrella of service design to better equip stakeholders, such as the company and consumer, in better visualising and articulating their specific needs and demands.

3.1 Tools available for Open Service Design

Service design is a useful framework for maintaining the perspective of the outside world of consumers as the leading element in selecting and elaborating ideas for potential innovation [13]. Service innovation managers have available a wide range of tools designed to engage various stakeholders in new service design. In the context of an open innovation approach to service design, companies attempt to engage 'outside in' innovation [10]; where outside stakeholders are actively brought into the design process [2].

In contrast to product design, where the end result is something tangible, a service is an intangible offering developed ultimately by a service innovation manager, an individual or a company [26]. This design process usually takes place behind closed doors, from service innovation managers working on their assessment of the consumers' requirements, traditionally informed by passive and reactive methods [18]. Increasingly, however, the lone ingenious designer, who could do everything by him/herself is rapidly becoming history [2] as it becomes more common for firms to search for knowledge existing beyond the boundaries of the company [24]. In line with increasingly complex customer demands, engaging with market-based partners such as customers and suppliers can help to better specify market requirements for specific services and to spread the costs and risks of the innovation process [24].

The following section discusses three potentially relevant design tools that are frequently used in service design. These are: **Lead User Research, Crowdsourcing and Prototyping**.

These tools have been selected on basis of their potential applicability to open service design. Academic research on service design is to date limited [1], especially so in the context of open innovation. The following section will hence discuss the possibilities and limitations of each of the selected open service design tools. One tool will then be selected that appears to best suit the open innovation approach to service design.

Lead User Research applied to Open Service Design. A term first coined by Eric von Hippel [16], lead user research focuses on individuals who are experiencing needs that are yet to be known to the wider public. These individuals, through intense use of certain company value propositions, often adapt existing value propositions to enhance the benefit they obtain from them. They typically have a strong intrinsic incentive to innovate on their own account [13] and are not necessarily bound by profits, organisational structures and production capabilities [12].

As stated by [13], integrating lead users into service design can be highly profitable and, in turn, can be valuable to involve lead users in the exploration stages of the innovation processes. Von Hippel [16] identifies lead users as possessing the following characteristics: (i) they face needs that will be in a general marketplace – but face them months or years before the bulk of that marketplace encounters them and (ii) they are positioned to benefit significantly by obtaining a solution to those needs. This principle follows the line of "necessity is the mother of invention", whereby lead users will engage in necessary problem solving to ensure that their needs are met. Market research with normal consumers is constrained by the familiarity of the product, its attributes, and the real-world experience to which the consumer has already been exposed. Therefore market opportunities, if any, uncovered by market research, represent incremental improvements in a product with reference to company competition rather than a consumer focus. However, since lead users' needs are often very specific, perhaps niche, adopting lead user research still leads to uncertainty when applying the service to a larger market.

Expecting success based on advanced and innovative users of a product, which is what lead users typically are, does not necessarily foreshadow general market adoption. Using lead users for open service design needs to therefore be approached with caution. Applying the needs of the few to the general community could result in a service being too complex to be feasible; straining a company to provide resources that could go beyond the needs of the wider community. Additionally, the specific consumers who can be lead users can be especially difficult to find. They may exist outside of targeted markets, be consumers of rival products, or, due to their rapidly evolving needs, could have already moved on to the next innovation by the time they are identified [16].

Crowdsourcing applied to Open Service Design. Although only officially recognised in 2005, engaging crowds in the generation of new ideas has been a useful open innovative tool used throughout history. The ever increasing availability of internet access and reach, however, has accelerated the range and complexity of consumer demands; making it more difficult for companies to address them behind closed company doors [32]. Companies are responding by giving more attention to gaining knowledge and insights from crowds, potential and actual consumers for example, as opposed to only from competitors [21]. The transition to 'outside-in' knowledge [10] goes against long-established approaches, in which companies typically considered their research and development efforts as the sole source of new technologies and products [12].

The utilisation of crowdsourcing within a company leverages economies of scale to solve problems that have been too resource-demanding to solve internally. Further, it is also used to generate ideas that will come closer to consumer expectations; creating, in turn, a lasting relationship between consumer and company as the opinions of consumers are listened to and acted upon. Companies engaging in crowdsourcing no longer see themselves only as resource providers, rather they are positioning themselves as community hubs - bringing together consumers and, for example, manufacturers and/or service providers [29].

The theoretically mass-directed approach that crowdsourcing permits does raise the question of product quality. Questions are also raised in response to its non-monetary incentive insofar that participants are less likely to commit for small, or no, financial incentive. Additionally, as the crowdsourcing participants are typically large in number, companies are required to engage resources (internal or outsourced) to direct, regulate and speak to participants, with a commensurate increase in management time and cost overheads [8]. While crowdsourcing may in principle work for new product development, services are intangible which involve processes and experiences that are difficult to define [18]. Although crowdsourcing is a useful tool for idea generation, its voluminous nature may complicate clear visualisations of services that provide value.

Prototyping applied to Open Service Design. Prototyping can be defined as the conceptualisation of abstract thoughts and specifications into tangible realisations. Where a prototype is the manifestation of ideas and the assumptions behind them, prototyping is an activity and a mindset [30] seeking to answer one or two questions at a time as opposed to the entire system [11]. Originally applied to manufacturing, prototyping in services faces intrinsic difficulties relating to a service's intangible nature. For this reason, prototyping in service design differs from its use in manufacturing in the sense of being a representation of a future situation [6].

Despite this apparent limitation, prototyping is a powerful service design tool that is able to identify current, future and perceived expectations of a service. Its nature is centred on participatory design and focuses largely on making ideas explicit by directly engaging stakeholders, particularly users of the service and others who will be impacted by the service, in the development of an effective design solution [17]. Engaging stakeholders in the development of services encourages multi-disciplinary input; removing the organisational barriers that can sometimes exist as remnants of the importance that corporate internal research and development once possessed [27]. The effect of prototyping is twofold: (i) from the management perspective companies gain inputs from divergent mindsets that would otherwise not have been discovered under traditional organisational structures (ii) by having participants demonstrate what a favourable service consists of they are breaking down the intangible complexities of a service. Such activities allow companies to address expectation gaps in turn minimising wastage of company resources. Giving permission to explore new behaviours - the tangible presence of a new thing, the prototype, itself encourages new behaviours, relieving individuals of the responsibility to consciously change what they do [11]. Hence the ideas generated are typically better representations of actual consumer demands and are better situated to address the improvement of a stakeholder experience.

While prototyping offers an empirical, user-centred, rapid solution, it is a practice still not easily understood nor applied in the service domain [30]. This is partly because services exist in stages with a multiplicity of touch points [37], thereby making it difficult to visualise and imagine how a service may be of benefit. Therefore the result of prototyping is constrained by the participants' ability to visualise the complexities of a service. The risk in this regard is the time and resources used to create a prototype [33] which may not be a real representation of the needs and wants of the wider community.

Flexibility within prototyping allows for this consumer involvement. However it does require that a company maintains open channels of communication between all stakeholders. Such an investment into new service design thinking requires a company to focus less on passive and reactive forms of research, rather more so on engaging with, listening to, and acting on stakeholder comments.

Summary: Crowd Sourcing, Lead User Research, & Prototyping.

	Possibilities	**Limitations**
Lead User Research	- Ahead of the mainstream market - Belong to a niche of innovators - Insights that may foreshadow future market trends	- Ahead of the mainstream market; assumed adoption - Potential overly complexity - Time expenditure with identifying lead users
Crowd-sourcing	- Company as a hub; pooling ideas from a large resource base - Outside-in approach breaking down company walls - Scale of knowledge reached far greater than internally possible	- Crowd dictates direction of the idea – possible misdirection - Difficulty in articulation with respect to service visualisations - Extra staff needed to manage incoming information flow
Prototyping	- Accurate visualisations of services. Breaking down services to a series of processes - Multidisciplinary input - An activity and mindset that is embodied in the thought process	- Dependent on participants ability to visualise - Multiplicity of touch points in services – difficultly in thinking broadly - Limited practical research

Figure 1: Identified possibilities/limitations of chosen service design tools

The importance of service design within the context of an increasing reliance on technology is indubitable. The preceding investigation into possibilities and limitations of the selected design tools, and their applicability to open service design, favoured prototyping as the most suitable in the context of co-creation and open service design. Prototyping's unique adaptability stems from the unique characteristic of stakeholder interaction and involvement; achieving a level of subconscious and 'in-use' reactions that the other tools lack. This real-life experience improves the speed at which consumers respond to stimuli in addition to providing visualisations of the experience that are closer to, if not the actual, demands of the consumer. The effect of which mitigates effects of misdirection and the saving of company resources [6]. Through participatory design companies switch their focus from providing a service they deem to be of value, rather allowing the consumer to tailor their own perspective of value. As ancillary developments in technology become more advanced, the integration of prototyping in conjunction with augmented and virtual reality will allow for scale and reach to broader markets.

Consumer interaction is typically characterised by passive and reactive methods to experiences [18], utilising retrospective descriptions to gain insight into consumer behaviours. Through interaction and involvement, prototyping in open service design allows companies to see how value is perceived from an outside-in perspective [10], transcending the intangibility of services. Although crowdsourcing is directed at large scale participatory involvement, it still exists within the context of passive and reactive interactions between company and customer. In this regard its capacity for participatory design are both the strength and weakness, in that it has the ability to engage scale, however only in a backdated capacity. Lead user research also encourages participation and involvement from consumers however lacks the scale of participants; assuming that the wider community will be receptive to a group of innovators.

The intangible characteristic of services contributes to the difficulty from which the visualising of services can occur. Prototyping engages stakeholders in intense visualisations, making ideas more tangible, complexity more readable and alternatives shareable [30]. Within service design it shows how different touchpoints, along with their associated actors and customer groups [37], fit together sequentially [30] making public services and hospitals, as examples, more accessible. Crowdsourcing lacks this ability to intensely visualise services as, due to its strength in scale, it can only exist in an online arena governed by moderators attempting to control the flow and direction of the innovative process. Consequently, and as Jonas et al. [18] argue, service industries are still lacking an open space for transparent, interactive value creation to include, on a large scale, customers, other potential future users, non-users and ordinary people. With lead user research, issues rise again regarding its adoptability from the general user community. As lead users' demands exist only within a select niche [13], it could be thought that integrating their ideas into open service design would serve to complicate services unnecessarily. Creating a service from such an advanced niche could create superfluousness, potentially straining company resources. In contrast prototyping in open service design is oriented around user-centred thinking. This implies that the walls between customer and company are, during prototyping in open service design, removed. The result of which is a collaboration of co-created effort leading to the realisation of a service that is both feasible and desired from the company and customer perspective respectively.

4 Executing Prototyping in Open Service Design

As services are experiences and a result of human-to-human interactions, service design is about communication [23]. In service innovation the challenge is to not only to identify the consumers' needs and wants they can express, but also to enable the consumers to tap into their subconscious desires [22] by instinctively interacting with their environment. Therefore, the role that the end-consumer plays is integral to prototyping within service design [22]. Dodgson [14] describes prototyping as the company's embodiment of open innovation and its direction towards the inclusion of an outside-in stream of knowledge. The following section will elaborate on the importance of prototyping and its methodologies to the service innovation manager [22]

4.1 How the Service Innovation Manager guides the process

Prototyping can be viewed as a tool that aligns the efforts of stakeholders such as companies or end-customers. Services are tested by having the users of the prototype interact in collaborative, explorative, iterative and open-ended ways. At the same time factors that could interfere with the service delivery and the user experience need to be taken into account as well [7]. Hence, service innovation managers can be seen as actors playing a key role as translators in the network of participating stakeholders. They ensure the user-centred design of the service by avoiding passive and reactive observations. Instead, they attempt to understand unrecognised needs of the customers and in turn propose relevant service solutions [4].

Service solutions offer value for the consumer in two ways: a customer evaluates a service (i) by its ability to fulfil a functional outcome and (ii) by its experiential offering [4]. As fulfilling functional targets is likely to lead to the consumer's satisfaction, providing an enjoyable experience during the process of a service may lead to a consumer's loyalty [4]. By prototyping the service experience, service innovation managers work towards meeting functional and experiential targets by proposing solutions and examining the results in a trial and error process; thereafter the result are refined in collaboration with the consumer. Buchenau & Fulton Suri [9] explain however that prototyping is not about using a toolkit or a set of techniques, rather about creating an attitude and language for communicating with the consumer in order to solve design problems.

4.2 Creating a Communication Tool Using Prototyping

The service design process is carried out in a number of phases [23]: Discovering, Concepting, Designing, Building and Implementing. Buchenau & Fulton Suri [9] write that prototyping is of value in understanding existing user experiences, exploring and evaluating design ideas, and communicating services to an audience. Participatory design, prototyping and co-creation are integral in the concepting, designing and building phases [23], and are ongoing activities throughout the design process [7]. Meanwhile, operating in the background, prototyping is functioning as a mechanism for knowledge transfer [23, 36]; hence breaking down the complexity of service systems into easier-to-visualise segments.

The consumer experiences a service as a journey in which interactions between consumer and service innovation manager are dynamic, complex, subjective, and go beyond concrete sensories [9]. Prototyping enables the identification and targeting of key touch points within this journey. Thus allowing both the service innovation manager and consumer to collaboratively break down a service experience into smaller segments. The cocreation thereafter is directed towards the improvement of smaller segments as opposed to the entire service system.

4.3 Leveraging Interaction and Involvement through Prototyping

Prototyping in service design provides the opportunity to describe, discuss and develop services efficiently. It achieves this, by (i) integrating customers as co-creators that can partake in creating services and (ii) contributing explicit and subconscious needs [18]. The different ways in which prototyping can be carried out are based on interactions between the users and emphasises giving voice to actors who may previously not have been involved in the innovation process. Through the process of doing and refining and by customer participation it is possible to create differentiated services [3].

4.4 Possible Caveats using Prototyping in Open Service Design

The inconsistent nature of services has implications on the prototype and the way people interact with it. The presentation of a technology-based prototype is the same each time, whereas a human-delivered prototype can vary even if the same people are involved [7]. This inconsistency could impact the validity of the produced prototype, hence risking the position of the service innovation manager, the collaborating consumers, and the relationship between them.

Awareness of fictitious scenarios could also influence the authenticity of the prototype. As people are privy to the role-playing scenario, there is a possibility that they behave differently in a real service situation. Additionally, it is possible that participants new to prototyping are unaware of relevant methodologies (e.g. theatre play, bodystorming, etc.). Hence resources of the service innovation manager may be required to train new participants and actively enourage open-minded participation in the service design process.

All the stakeholders of a service should ideally be present already at the prototyping stage in order to improve how similar test and implementation contexts are [7]. As the service experience is ultimately subjective, producing only one prototype would not be considered sufficient. Therefore, several prototypes would be necessary to validate the consistency and authenticity of the service design.

4.5 Open Service Design in Practice

(Open) service innovation as a research field is a relatively young topic that is receiving increasing traction. Meanwhile, an increasing number of 'service innovation labs' (c.f. Fraunhofer ServLab[1], JOSEPHS®: The Service Manufactory[2] or the Service Innovation Corner[3] among others) are interesting approaches where prototyping methodologies in the development of new service concepts are applied.

Service innovation labs equipped with sophisticated technology enable abstract ideas to transcend into a tangible reality. Positioned as an engagement platform between service innovation manager and consumer, they apply the principles of prototyping through active participation and collaboration. This in turn encourages and promotes cocreation of services from different stakeholders early in the innovation process.

The Service Innovation Corner – University of Lapland, Finland. The SINCO lab is described as a technology-aided prototyping environment for user experience in which tools and devices support the concretisation of ideas, testing and agile co-creation [33]. It is an ideal example of how prototyping as a service design tool can approach the aforementioned challenges of services. The focal point of this approach is to map the divergent touchpoints in a consumer's experience of a service. Technology is hence used as an aid to create a 'transferable model' of a service prototyping environment [23].

SINCO identifies factors that need to be considered during prototyping. Thus the list below can be seen as a guideline for service innovation managers:

- Understanding challenges: users, environment and technologies involved,
- Creating empathy for and co-operating with the users,
- Including clients, other stakeholders and customers in the process,
- Prototyping, improving and visualising methods during the whole process and,
- Implementing, maintaining and continuous development at later stages.

The service labs extensive technological capabilities enable the creation of a 'Servicescape[4]', thus simulating the real service experience to help overcome authenticity and validity difficulties. Additionally, a skillful group leader and technical expert are necessary to include new participants and real consumers who are unfamiliar

---

[1] Fraunhofer-IAO: http://www.servlab.eu/
[2] JOSEPHS® - The Service Manufactory: http://www.josephs-service-manufaktur.de/
[3] The Service Innovation Corner (SINCO), University of Lapland: http://sinco.fi/
[4] An artificial and physical environment of the place in which a service is provided (Bae & Leem, 2014)

with the concept and enable them to freely immerse themselves into a creative state that is required for prototyping [36]. Through the combination of new technology with prototyping in the SINCO lab, agile collaboration with users is easily achieved [22].

JOSEPHS® - The Service Manufactory. JOSEPHS® - The Service Manufactory, located in the heart of Nuremberg, Germany, is a non-profit innovation laboratory and an innovation ecosystem for the development and testing of novel service concepts and prototypes. The goal of the experimental initiative is to provide an intermediary platform for consumers and companies to actively engage in co-creative and collaborative innovations [34]. This is achieved by means of alternating exhibitions, or theme worlds, in which members of the public have a chance to view, interact with, and leave feedback for companies displaying their novel service or product prototypes. This offers users a rare glimpse into the thinking behind the company and thus allows users to play a (pivotal) part in the early stages of product or service development.

Opening up the prototyping process to the public means that companies operating through JOSEPHS® are able leverage participatory scale. Moreover, through the regular hosting of diverse events in the Denkfabrik (Thinking Factory), JOSEPHS® offers a platform for forward-thinking innovators as well as everyday consumers to collaborate with companies. Furthermore, companies have the chance to enact agile-like methods on their products and services based on the nature of feedback received. Hence the product or service remains in a dynamic state until the end of a company's three-month tenure at JOSEPHS®. This in turn allows for the development of a valid and authentic product or service before launching into broader environments.

5 Future Outlook: Prototyping in Virtual Reality

In the past, 'virtual worlds' or 'computer-simulated realities' were mainly used for public entertainment, cultural projects, virtual museums etc. [3]. Advances in computer hardware and gaming technology in the last decade have given rise to a new generation of virtual world applications running on desktop computers [19]. This new dimension adds another way in which service innovation managers and consumers can practice prototyping in open service design. Consequently, the application of collaborative virtual environments (CVEs) [38] in service-oriented fields is increasing and becoming more diverse. By means of high-immersion displays, haptic virtual environments, and other sensory-devices [28], virtual worlds are able to provide exceptional depth in a collaboration-driven prototyping experience. This is in contrast to simple online environments which often cannot provide the suitable relational depth of experiences and emotions needed for constructive co-creation [18].

5.1 Application and Relevance to Open Service Design

From a practical standpoint, virtual reality prototypes can, due to their novel use in the service industry, be more time consuming and costly compared to traditional prototyping tools. Additionally, users, just as with programmers designing service platforms, require new mindsets to become part of a structure in a virtual service organisation [15]. Therefore, the realistic presentation of a service scape and concept demands high effort and resources for scenery composition and programming from the service provider's side and an affinity for virtual worlds from the user's side [31]. Regardless, the hitherto status quo for gaining consumer knowledge is through market research, whereby statistical studies are regarded as being more scientific. Yet, it is known that asking the customer, particularly through structured questionnaires, only reveals a superficial layer of attitudes and behaviour, rather than root motivations [15]. In contrast, virtual worlds are inherently interactive, allowing users to interact with simulated artefacts [20] as they would in real settings. Thus this process of interaction creates possibilities for service innovation managers to easily integrate users in the development process of a service [31].

Virtual worlds are powerful platforms for designing novel collaborative design environments [20] and have been used frequently in architecture and organisational and interior design spaces [25]. Furthermore, virtual worlds uniquely allow users the embodiment of another character through avatars. The added persona enables users to be more than just viewers of a virtual space. Rather, through their observable motions and actions they are able to express both verbal and non-verbal forms of communication and awareness. According to Koutsabasis [20], "the avatars' position and orientation communicate where they are and what are they looking at; their appearance can usually be modified to express the user's personality, or even to denote the role of the user in a collaborating team; their animated bodies communicate their current activity; in some virtual worlds the avatars may also use facial expressions and gestures as an additional means of communication. All these abilities are important for the quality of the remote communication and coordination of a design team." Thus, virtual worlds provide an interactive platform for service innovation managers and users in which prototyping can be iteratively practiced producing a refined service innovation.

5.2 Adaptability of Virtual Worlds to Subjective Service Experiences

Virtual worlds enforce the characteristics of prototyping without having to rely on a user's physical presence. Therefore, by eliminating mitigating factors such as presence, weather, or proximity, a whole new service design paradigm could be conceptualised from the comfort of one's own home. As service innovation managers would be able to record and analyse more than just traditional and retrospective market research, they would be confronting and overcoming the intangible and experience-driven characteristics of a service.

Moreover, virtual worlds offer their own scripting language to be used for extending the real world's functionality and defining specific object behaviour. This makes them customisable and allows developers to design and implement application-specific tools whilst taking advantage of the existing visualisation, interaction and communication infrastructure of the virtual worlds [20]. Therefore programmers and service innovation managers with the suitable skill set can deconstruct, reuse, and rebuild platforms based on the size of their existing graphic and source code libraries. In addition to providing numerous cost efficiencies, this would be particularly useful for a service innovation manager working across a number of industries.

6 Conclusion and Postulations for Future Research Directions

The importance of service design in a service-dominant economy is undeniable. With the rapid developments of, and reliance on, information technology, the service innovation manager is faced with a difficult role in catering for increasingly complex demands of consumers.

Open service design is an open innovation technique that encourages non-traditional involvement from outside stakeholders. Through the analysis of three different service design tools it was suggested that prototyping was not only the most appropriate tool in the service designer's arsenal, but matched open service design as an ideology.

Given the intangible and complex nature of services, prototyping still faces challenges of validity and authenticity, particularly with unfamiliar participants. Institutions such as SINCO and JOSEPHS® are approaching these challenges by encouraging and fostering an outside-in stream of knowledge. Both institutions display interesting and innovative methods in which service innovation managers can test the validity and authenticity of a service idea.

Integrating prototyping into future technologies such as virtual worlds have the potential to alter methodologies in service design; potentially reaching a larger audience from which to apply prototyping iteratively. The inclusion of novel technology has been designed to provide a real world scenario where participants can actively participate without having to leave the comfort of their own homes. Looking to virtual worlds may therefore answer the question as to how services, with their characteristic multiplicity of touch points, can be fully visualised.

In lieu of the wide-scale availability of virtual worlds, it could be suggested that more documentation and research is needed into the processes of prototyping in open service design. Namely, at which particular stage of open service design is the prototyping process most vulnerable to inauthenticity and lack of validity? Or, what contextual factors are companies and consumers most subject to in order to collaboratively prototype a service that can benefit both parties? Literature in this direction would develop further the field of research in service innovation, prototyping and open innovation. Additionally, it would ultimately help service innovation managers engage and collaborate more effectively with their consumers.

References

1. Akama, Y.: Warts-and-all: the real practice of service design. First Nord. Conf. Serv. Des. Serv. Innov. Oslo 24th-26th Nov 2009. 61, 0, 1–11 (2009).
2. Albinsson, L. et al.: Towards a Co-Design Approach for Open Innovation. Des. Co-designers Work. 1, 5, 1–5 (2008).
3. Bae, Dae Jung, Leem, C.S.: Managing Service Quality Article information : (2014).
4. Beltagui, A. et al.: Design in the Experience Economy: Using Emotional Design for Service Innovation. Adv. Int. Mark. 23, 195–205 (2012).
5. Bhat, W.A.Q.S.M..: Industrial Management & Data Systems. Access. (2014).
6. Blomkvist, J.: Representing Future Situations of Service. Prototyping in Service Design. (2014).
7. Blomkvist, J., Holmlid, S.: Service designers on including stakeholders in service prototyping. Proc. 6th Int. Conf. Incl. Des. – Incl. 2011. 1–10 (2011).
8. Borst, I.: The Case for and against crowdsourcing (part 2). (2015).
9. Buchenau, M. et al.: Experience Prototyping. Conf. Des. Interact. Syst. Process. Pract. methods, Tech. 424–433 (2000).
10. Chesbrough, H.: Bringing Open Innovation to Services. Mit Sloan Manag. Rev. 52, 2, 85+ (2011).
11. Coughlan, P. et al.: Prototypes as (Design) Tools for Behavioral and Organizational Change: A Design-Based Approach to Help Organizations Change Work Behaviors. J. Appl. Behav. Sci. 43, 1, 122–134 (2007).
12. Delden, C. van: The Power of Users as Innovators. In: Crowdsourced Innovation: Revolutionising Open Innovation with Crowdsourcing. pp. 37–44 Innosabi GmbH, Munich (2014).
13. Dijk, G. Van, Raijmakers, B.: Open Innovation as a service design approach. (2009).
14. Dodgson, M. et al.: The role of technology in the shift towards open innovation : the case of Procter & Gamble. R&D Manag. 36, 3, 333–346 (2006).

15. Gummesson, E.: Service Management : An Evaluation and the Future. Int. J. Serv. Ind. 5, 1, 77–96 (2006).
16. Hippel, E. von: Lead Users: An Important Source of Novel Product Concepts. Manage. Sci. 1, 7, 791–805 (1986).
17. Holtzblatt, K., & Jones, S.: Contextual inquiry: A participatory technique for systems design. Erlbaum, Lawrence, Hillsdale (1993).
18. Jonas, J. et al.: Open Service Design ? Exploring Customer Co-creation in a Service Manufactory. 477–480 (2014).
19. Kan, H.Y. et al.: An Internet virtual reality collaborative environment for effective product design. Comput. Ind. 45, 2, 197–213 (2001).
20. Koutsabasis, P. et al.: On the value of Virtual Worlds for collaborative design. Des. Stud. 33, 4, 357–390 (2012).
21. Mention, A.-L.: Co-operation and co-opetition as open innovation practices in the service sector: Which influence on innovation novelty? Technovation. 31, 1, 44–53 (2011).
22. Miettinen, S. et al.: Experience Design in Digital Services. Res. Econ. Bus. Cent. East. Eur. 6, 1, 29–50 (2014).
23. Miettinen, S. et al.: Realizing Design Thinking through a Service Design Process and an Innovative Prototyping Laboratory–Introducing Service Innovation Corner (SINCO). Sinco.Fi. July, 1–4 (2012).
24. Mina, A. et al.: Open service innovation and the firm’s search for external knowledge. Res. Policy. 43, 5, 853–866 (2014).
25. Mobach, M.P.: Do virtual worlds create better real worlds? Virtual Real. 12, 3, 163–179 (2008).
26. Morelli, N.: Designing Product / Service Systems: A Methodological Exploration. Des. Issues. 18, 3, 3–17 (2002).
27. Möslein, K.M.: The Emergence of Platforms for Open and Crowdsourced Innovation. In: Crowdsourced Innovation: Revolutionising Open Innovation with Crowdsourcing2. pp. 60–70 Innosabi GmbH, Munich (2014).
28. MUÑOZ, J.M.: A Vibrotactile Prototyping Toolkit for Virtual Reality and Videogames. 28–41 (2011).
29. Niklas, A.: Branded Communities for Crowdsourced Innovation. In: Crowdsourced Innovation: Revolutionising Open Innovation with Crowdsourcing. pp. 46–58 Innosabi GmbH, Munich (2014).
30. Passera, S. et al.: When, how, why prototyping? A practical framework for service development. Brown 2008, 1–16 (2012).
31. Rau, C. et al.: Open Tourism. 171–187 (2016).
32. Riemensberger, F.: The importance of Crowsourcing. In: GmbH, I. (ed.) Crowdsourced Innovation: Revolutionising Open Innovation with Crowdsourcing. pp. 12–16 , Munich (2014).
33. Rontti, S., Lindström, A.: Tools and Methods for Technology-Aided Prototyping of User Experience: SINCO Environment as a Pilot of User Experience. (2014).
34. Roth, A. et al.: Interaktive Kunden als Herausforderung: Die Fallstudie „JOSEPHS® – Die Service-Manufaktur“. HMD Prax. der Wirtschaftsinformatik. 51, 6, 883–895 (2014).
35. Siggelkow, N.: Persuasion with Case Studies - Siggelkow - S1&2 R1.pdf. 50, 1, 20–24 (2007).
36. Simo, R. et al.: A Laboratory Concept for Service Prototyping – Service Innovation Corner (SINCO). ServDes. Serv. Des. Innov. Conf. 8-10 Febr. 2012, Helsinki. 229–241 (2012).
37. Stickdorn, M., Schneider, J.: This is Service Design Thinking. BIS Publishers, Amsterdam (2010).
38. Vosinakis, S. et al.: Virtual Environments for Collaborative Design : Requirements and Guidelines from a Social Action Perspective. X, X, (2006).
39. Zeithaml, V. et al.: Delivering Quality Service: Balancing Customer Perceptions and Expectations. J. Econ. Lit. 226 (2007).

Zielgruppendilemma des gleichzeitigen stationären und Online-Handels: Eine experimentelle Studie am Beispiel des Facebook-Auftritts eines Mode-Einzelhandelsunternehmens

Mario Porst, Sven Dittes, und Stefan Smolnik

FernUniversität in Hagen, Hagen, Deutschland
mporst22@gmail.com,{sven.dittes,stefan.smolnik}@fernuni-hagen.de

Abstract. Immer mehr traditionelle Einzelhandelsunternehmen sehen sich mit den Auswirkungen der Digitalisierung konfrontiert und eröffnen eigene Online-Shops, um mit E-Commerce-Plattformen wie Amazon und Zalando mithalten zu können. Hierbei setzen sie zumeist auf die über Jahrzehnte stationär aufgebaute Marke, unter der sie dann auch den Online-Shop führen. Darüber hinaus haben diese Unternehmen auch Social-Media-Plattformen wie Facebook als kosteneffizienten Marketingkanal für sich entdeckt. Da über diese Kanäle jedoch sowohl die Kunden des stationären Kaufhauses als auch die des Online-Handels erreicht werden, stellt sich vielen Unternehmen die Frage, wie sie mit demselben Online-Auftritt beide Zielgruppen adressieren können. In diesem Beitrag wird eine Fallstudie vorgestellt, mit deren Hilfe dieses Zielgruppendilemma untersucht wird. Darauf aufbauend wird eine experimentelle Studie konzipiert, um festzustellen, ob Beiträge auf sozialen Medien von beiden Zielgruppen unterschiedlich bewertet werden. Die Ergebnisse können von Unternehmen verwendet werden, um ähnliche Zielgruppendilemmata zu adressieren und die eigenen Aktivitäten auf Social-Media-Plattformen strategisch und fundiert zu planen.

Keywords: Social Media, stationärer Handel, Online-Handel, E-Commerce, Facebook

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Porst, M.; Dittes, S.; Smolnik, S. (2017): Zielgruppendilemma des gleichzeitigen stationären und Online-Handels: Eine experimentelle Studie am Beispiel des Facebook-Auftritts eines Mode-Einzelhandelsunternehmens, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1441-1452

1 Einleitung

Social-Media-Plattformen (SMP) wie Facebook und Twitter, aber auch neuere SMP wie Pinterest und Instagram haben sich im privaten Leben zum Zweck des Informationsaustauschs und der Vernetzung mit Freunden etabliert [1-5]. Die sich daraus ergebenden Möglichkeiten haben auch Unternehmen erkannt, die zunehmend auf diesen Plattformen mit eigenen Profilen und Auftritten vertreten sind. So stieg der Prozentsatz der Unternehmen, die SMP einsetzen, von 42% im Jahr 2008 auf 88% im Jahr 2012 an [6]. Nachdem Social Media bis vor einigen Jahren nur als „Marketinghype" von einigen Online-Spezialisten betrachtet wurde, ist es heute ein akzeptierter Marketingkanal [7]. Laut einer Studie von Bruhn, Martin und Schnebelen [8] schreibt mehr als die Hälfte der befragten deutschen Unternehmen sozialen Medien einen relativ hohen oder sehr hohen Stellenwert zu.

Das B2C-E-Commerce-Wachstum setzt sich laut dem Handelsverband Deutschland (HDE) weiter fort. Dieser prognostiziert für das Jahr 2016 einen Umsatz ohne Umsatzsteuer von 46,3 Milliarden Euro. Dies entspricht einem Wachstum von fast 300% in den letzten zehn Jahren[1] [9]. In einer weiteren Studie des HDE wurden Händler befragt, ob sie bereits einen Online-Shop haben oder planen, in den nächsten zwölf Monaten einen Online-Shop zu eröffnen. Fast jeder dritte (28,9%) stationäre Händler hatte in der Befragung bereits einen Online-Shop aufgebaut. Weitere 8,8% planen die Eröffnung in den nächsten zwölf Monaten [10]. Ein Vorteil von E-Commerce gegenüber dem stationären Handel ist, dass das Einzugsgebiet vergrößert wird, da Online-Shops über das Internet überregional erreichbar sind. Für Unternehmen mit regional fokussiertem stationärem Handel bedeutet dies eine Ausweitung des Einzugsgebiets.

Dieses Potenzial hat auch ein Mode-Einzelhandelsunternehmen erkannt, das im weiteren Verlauf des Beitrags als Fallstudienunternehmen dient. Dieses stationär regional fokussierte Unternehmen hat 2006 einen Online-Shop eingerichtet, um seine Modeartikel auch überregional verkaufen zu können. Jedoch sind mit dieser Ausweitung auch Herausforderungen verbunden. Beispielsweise ist die Zielgruppenansprache auf SMP nicht trennscharf umzusetzen. Wenn z.B. die regionalen Fans der entsprechenden Facebook-Seite des Unternehmens in der Regel eher an Informationen über das Unternehmen und dem stationären Ladengeschäft interessiert sind, sind – aufgrund der überregionalen Reichweite des Online-Shops – die Fans, die nicht im stationären Einzugsgebiet leben, eher an Angeboten und Informationen, die den Online-Shop betreffen, interessiert. So entstand für das Fallstudienunternehmen ein Zielgruppendilemma auf diesen Plattformen: Entweder werden die regionalen Kunden zufriedengestellt und die überregionalen Kunden mit nicht relevanten Informationen, wie z.B. Aktionen im Ladengeschäft, konfrontiert oder es werden den regional interessierten Kunden zu wenig Informationen über das stationäre Ladengeschäft mitgeteilt. Facebook-Seiten sollten ihren Fans jedoch interessante, unterhaltsame und innovative Inhalte liefern [11]. Dieser Beitrag adressiert dieses Zielgruppendilemma des Fallstudienunternehmens und hat das Ziel,

[1] 15,7 Milliarden Euro ohne Umsatzsteuer im Jahr 2006

eine wissenschaftliche Grundlage zu erarbeiten und eine Lösung zu finden, wie mit diesem Problem umgegangen werden kann.

Im Folgenden wird zunächst das Fallstudienunternehmen vorgestellt, das sich gerade in einem Zielgruppendilemma hinsichtlich seiner Marketingstrategie für Facebook befindet. Aufbauend auf den Erkenntnissen aus dem Fallstudienunternehmen werden in diesem Beitrag die Ergebnisse einer experimentellen Studie vorgestellt. Aufgeteilt in zwei Gruppen wurde eine Umfrage bei Facebook-Nutzern durchgeführt, bei der die Teilnehmenden zunächst jeweils in ein unterschiedliches Szenario eingeführt wurden. In einem Szenario wurde den Teilnehmenden eine Facebook-Seite eines Mode-Einzelhandelsunternehmen vorgestellt, das sie nur als Ladengeschäft kennen, und in einem anderen Szenario nur als Online-Shop. Beiden Gruppen wurden Bilder derselben Facebook-Beiträge eines fiktiven Unternehmens vorgelegt, die schließlich von den Teilnehmenden bewertet wurden. Als Ergebnis der Studie werden Unterschiede hinsichtlich der Bewertung der Facebook-Beiträge präsentiert. Auf Basis dieser Ergebnisse werden Hypothesen bezüglich der festgestellten Unterschiede formuliert und Handlungsempfehlungen für Unternehmen abgeleitet.

2 Theoretische Grundlagen

2.1 Facebook-Seiten und deren Möglichkeiten

Mit einer Facebook-Seite können Unternehmen ihre Kunden und Interessierte mit verschiedenen Arten von Informationen versorgen. Sie haben die Möglichkeit, die Kunden und Interessierten selektiv, sogar auf personalisierter Ebene, über ihre Produkte und Dienstleistungen zu informieren, sie mit interessanten oder nützlichen Inhalten zu erreichen oder die Facebook-Seite als einen Kundenservice-Kanal zu nutzen [12]. Unternehmen können auf ihren Facebook-Seiten Beiträge veröffentlichen, die Anekdoten, Fotos, Videos, Veranstaltungen, Hyperlinks oder anderes Material beinhalten. Ihre Fans können mit diesen Beiträgen dann interagieren, indem sie auf „Gefällt mir“ klicken oder die Beiträge kommentieren sowie teilen [13]. Weiterhin gibt es eine Reihe von zusätzlichen Möglichkeiten für Unternehmen wie z.B. das Schalten von Werbeanzeigen.

Grundsätzlich können auf die sozialen Medien klassische Marketingprinzipien angewendet werden [14]. So besteht die Aufgabe von Werbemitteln darin, dass die vom Unternehmen gewollten Botschafts- bzw. Beitragsinhalte mit der gewünschten Wirkung an den Nutzer (Rezipient) herangetragen werden [15]. Daher müssen auf SMP Inhalte geliefert werden, die den Nutzern (Fans) angenehme Erfahrungen bieten, sodass diese wiederkommen und aktiv mitwirken [16]. Hierzu schreiben Castronovo und Huang [16], dass eine virale Kampagne eher erfolgreich ist, wenn der Inhalt der Nachricht stark auf die Zielgruppe abgestimmt ist und somit auf die Motivation zum Teilen dieser Informationen einwirkt. Des Weiteren stellen Heller-Baird und Parasnis [17] fest, dass Unternehmen Erlebnisse bereitstellen müssen, die den Kunden und Interessierten einen konkreten Mehrwert im Gegenzug zu Zeit, Aufmerksamkeit, Unterstützung und Daten bieten, um das Potenzial von Social Media erfolgreich zu nutzen.

2.2 Wahrnehmung, Einstellung und Reaktion auf Facebook-Beiträge

Um messen zu können, wie ein Beitrag auf die Rezipienten wirkt und ob dieser einen Mehrwert liefert, wird im Folgenden auf das Konzept des (Botschafts-) Involvements zurückgegriffen. Trommsdorf [18] definiert Involvement als den „Aktivierungsgrad bzw. die Motivstärke zur objektgerichteten Informationssuche, -aufnahme, -verarbeitung und -speicherung“. Weiterhin unterscheidet Trommsdorff diverse Kategorien von Involvement. Ob eine Botschaft subjektiv mehr oder weniger interessant ist, beschreibt dabei – unabhängig vom Produkt, für das die Werbung gemacht wurde – der Grad des Botschaftsinvolvements [18]. Im Folgenden soll der Begriff Involvement für das Botschaftsinvolvement stehen und damit beschreiben, ob die Facebook-Beiträge des Untersuchungsobjekts für die Betrachter subjektiv mehr oder weniger interessant sind.

Um das Involvement auf verschiedenen Ebenen abzufragen, wurden die drei Bereiche *Wahrnehmung*, *Einstellung* sowie *Reaktion* festgelegt. Die *Wahrnehmung* (kognitive Komponente) ergibt eine *Einstellung* (affektive Komponente), die wiederum zu einer *Reaktion* (konative Komponente) führt. Um die Wahrnehmung und den Wert von Werbung im Internet zu messen, hat Ducoffe [19] verschiedene Dimensionen definiert: *Informationsgehalt*, *Unterhaltung*, *Verärgerung* und *Werbewert*. Weiterhin wird die *Einstellung* zu den Beiträgen gemessen. Innerhalb des Bereiches *Reaktion* wird die Intention zu jeweils „Gefällt mir“, „kommentieren“, „teilen“, „klick auf Link“ oder „Interessiert“ abgefragt. Reaktionen auf Beiträge werden von Werbetreibenden und Marketer als Metriken für die Erfolgsmessung von Kampagnen herangezogen [20]. Im Rahmen der nachfolgend beschriebenen Fallstudie werden diese Dimensionen innerhalb verschiedener Gruppen von Rezipienten gemessen, analysiert und miteinander verglichen.

3 Fallstudie

Das Mode-Einzelhandelsunternehmen ist seit über 100 Jahren in einer deutschen Metropolregion ansässig und stark mit der Region verwachsen. Seit der Einrichtung des Online-Shops wird über die Region hinaus Ware angeboten. Die Untersuchungen innerhalb dieser Studie beziehen sich auf die 2010 erstellte Facebook-Seite des Unternehmens. Am 10.02.2016 weist diese ca. 90.000 „Gefällt mir“-Angaben auf. Die Geschlechterverteilung der Fans der Seite ist laut Facebook-Seitenstatistik am 06.03.2016 mit 83% Frauen und 17% Männern angegeben. Mit 41% hat die Gruppe der 18- bis 24-Jährigen den größten Anteil der Fans, gefolgt von den 13- bis 17-Jährigen mit 27% und den 25- bis 34-Jährigen mit 20%.

Die Facebook-Seite wird derzeit als relevanter Marketingkanal für Kunden und Interessierte sowohl für den überregionalen E-Commerce als auch für den regionalen stationären Handel genutzt. Aus diesem Grund befindet sich das Unternehmen in einem Zielgruppendilemma und stellt sich u.a. folgende Fragen:

- Wie und in welchem Maße sollten beide Kundengruppen angesprochen werden?
- Werden Beiträge mit E-Commerce-Inhalten von der Zielgruppe der Kunden und Interessierten des regionalen stationären Handels schlechter bewertet?
- Werden Beiträge mit Ankündigungen zu stationären Aktionen von der Zielgruppe des überregionalen E-Commerce schlechter bewertet?
- Werden neutrale Inhalte, wie z.B. zu Trends und Styling-Tipps sowie Entertainment, von beiden Zielgruppen gleich bewertet?

Diese Studie soll ein besseres Verständnis über das Zielgruppendilemma schaffen und damit dazu beitragen, Lösungsansätze zu finden, wie dieses Problem zu adressieren ist. Dazu konnten im Vorfeld gemeinsam durch Interviews mit relevanten Stakeholdern des Fallstudienunternehmens folgende Ziele identifiziert werden:

- Reduktion von Kosten bzw. effizientere Nutzung von vorhandenen Ressourcen
- Beide Zielgruppen gleichermaßen erreichen
- Beide Zielgruppen zielgruppengerecht ansprechen
- Strategisch ausgerichtete Social-Media-Aktivitäten auf Basis von wissenschaftlicher Forschung

4 Methodisches Vorgehen

4.1 Voruntersuchung

Um den Status-quo der Aktivitäten auf der untersuchten Facebook-Seite zu extrahieren, wurde deren Inhalt analysiert. Hierfür wurden alle Beiträge im Zeitraum vom 01.11.2015 bis zum 31.01.2016 betrachtet und klassifiziert. Die Zahl der analysierten Beiträge in diesem Zeitraum beträgt 308. Dabei wurden die Beiträge nach dem *Inhalt*, den ausgewählten *Typen* von Beiträgen und der *Ausrichtung* kategorisiert. Tabelle 1. zeigt sechs Kombinationen der identifizierten Beitragscluster.

Tabelle 1. Ausgewählte Kombinationen von Beiträgen

	Inhalt				*Typ*			*Ausrichtung*	
	E-Commerce	*Trends & Styling-Tipps*	*Stationäre Aktion*	*Enter-tainment*	*Bild*	*Link*	*Veran-staltung*	*regional*	*über-regional*
Beitrag #1	x				x				x
Beitrag #2	x					x			x
Beitrag #3		x				x		x	x
Beitrag #4			x		x			x	
Beitrag #5			x				x	x	
Beitrag #6				x	x			x	x

Auf Basis dieser Voruntersuchung wurden sechs entsprechende Beiträge aus dem analysierten Zeitraum ausgewählt, die daraufhin in der Fragebogenstudie verwendet wurden. Im folgenden Kapitel wird das Vorgehen bei der experimentellen Fragebogenstudie beschrieben.

4.2 Experimentelle Fragebogenstudie

Im Rahmen der Studie wurden in einem Online-Fragebogen den Teilnehmern Bilder von Facebook-Beiträgen gezeigt. Die Beiträge wurden auf Basis der Analyse beim Fallstudienunternehmen ausgewählt (siehe Tabelle 1). Für die Beiträge #2 und #3 wurden, da es sich um Beiträge mit Inhalten zu Produkten, Styling-Tipps und Trends handelte, jeweils eine Version für weibliche und männliche Probanden genutzt und die Ergebnisse aggregiert.

Alle sechs Beiträge wurden mithilfe von 5-Punkt-Likert-Skalen von den Teilnehmenden bewertet. Es wurden Konstrukte für die drei Bereiche *Wahrnehmung*, *Einstellung* sowie *Reaktion* untersucht (siehe Abschnitt 2.2). Tabelle 2 zeigt eine Übersicht über die in der Studie erhobenen Konstrukte sowie die Items und die Quellen, aus denen die Items abgeleitet wurden.

Tabelle 2. Übersicht der verwendeten Konstrukte und Items

Bereich	*Konstrukt*	*Item*	*Quelle*
Wahrnehmung	Informationsgehalt	Dieser Facebook-Beitrag ist eine gute Informationsquelle.	Brackett und Carr [21] Ducoffe [19]
		Dieser Facebook-Beitrag gibt mir relevante Informationen.	
		Dieser Facebook-Beitrag beinhaltet aktuelle Informationen.	
	Unterhaltung	Dieser Facebook-Beitrag ist unterhaltsam.	Brackett und Carr [21] Ducoffe [19]
		Dieser Facebook-Beitrag ist angenehm.	
		Dieser Facebook-Beitrag ist ansprechend.	
	Verärgerung	Dieser Facebook-Beitrag ist störend.	Brackett und Carr [21] Ducoffe [19]
		Dieser Facebook-Beitrag ist lästig.	
		Dieser Facebook-Beitrag ist irritierend.	
	Wert des Beitrags	Dieser Facebook-Beitrag ist nützlich.	Brackett und Carr [21] Ducoffe [19]
		Dieser Facebook-Beitrag ist wertig.	
		Dieser Facebook-Beitrag ist wichtig.	
Einstellung	Einstellung zu dem Beitrag	Ich finde diesen Facebook-Beitrag gut.	Burton und Lichtenstein [22]
Reaktion	Intention „Gefällt mir“	Ich würde „Gefällt mir“ bei diesem Beitrag klicken.	Alhabash et al. [20]
	Intention „kommentieren“	Ich würde diesen Beitrag kommentieren.	
	Intention „teilen“	Ich würde diesen Beitrag „teilen“.	
	Intention „Interessiert“ bzw. klick auf Link	Ich würde auf „Interessiert“ bzw. den Link klicken um weitere Informationen zu erhalten.	

Die Fragebogenstudie wurde in zwei Schritten durchgeführt. Im ersten Schritt wurde die Umfrage bei den Fans des Mode-Einzelhandelsunternehmens der Fallstudie durchgeführt (im Folgenden als „echte“ Umfrage betitelt). Die Beiträge waren hier nicht verfremdet und mit dem Firmenlogo (Profilbild) versehen.

Im zweiten Schritt wurde der experimentelle Teil der Studie durchgeführt, indem die Teilnehmenden zunächst in das jeweilige fiktive Unternehmen „MODE haus“ eingeführt wurden. Die Facebook-Beiträge wurden hierfür anonymisiert umgesetzt. So wurde sichergestellt, dass eine Beeinflussung durch die Marke auszuschließen ist. Die Teilnehmenden wurden randomisiert auf zwei verschiedene Gruppen aufgeteilt. Zu Beginn des Fragebogens wurden die Teilnehmenden jeweils in ein Szenario eingeführt: Das Szenario 1 suggerierte, dass die Teilnehmenden das Unternehmen nur als stationäres Geschäft kennen und noch nie online bei dem dazugehörigen Online-Shop eingekauft haben. Szenario 2 suggerierte den Teilnehmern, dass sie noch nie vor Ort waren und keinen Bezug zum stationären Ladengeschäft haben, jedoch bereits beim dazugehörigen Online-Shop eingekauft haben. Bei allen drei Versionen der Umfrage wurden die gleichen Beiträge verwendet.

Der experimentelle Teil der Fragebogenstudie wurde im Zeitraum vom 08.05.2016 bis zum 23.05.2016 durchgeführt. Der entsprechende Fragebogen wurde am 08.05.2016 mittels eines Facebook-Beitrages über private Facebook-Profile veröffentlicht, sowie am 10.05.2016 über den Verteiler der Stiftung Begabtenförderung berufliche Bildung gGmbH (SBB) an 4227 E-Mail-Adressen versendet. Am 19.05.2016 wurde über den SBB-Verteiler eine Erinnerung gesendet. Das Online-Befragungswerkzeug wies den Teilnehmenden randomisiert eines der beiden Szenarien zu. Die „echte“ Umfrage war nicht Bestandteil der Randomisierung und wurde daher über einen separaten Hyperlink am 10.05.2016 mittels einer Facebook-Anzeige über das Unternehmensprofil des Fallstudienunternehmens veröffentlicht. Um die Beteiligungsquote zu erhöhen, wurden bei allen Umfragen Anreize in Form von Verlosungen von Sachgeschenken und Gutscheinen eingesetzt.

5 Ergebnisse

Von 580 abgeschlossenen Umfragen der Fragebogenstudie haben 489 Teilnehmende angegeben, dass sie ein Facebook-Profil für private Zwecke besitzen. Von diesen 489 hatten 414 Teilnehmende ein Aufmerksamkeits-Item korrekt beantwortet. Die nachfolgenden Ergebnisse stellen daher nur die Ergebnisse dieser 414 Teilnehmenden dar. In der Abbildung 1 werden die Ergebnisse der erhobenen demografischen Merkmale dargestellt. 75,85% (n=314) der Teilnehmenden waren weiblich und 24,15% (n=100) waren männlich. Auf die „echte“ Umfrage entfielen 129 ausgefüllte Fragebögen, auf die Umfrage des Szenarios 1 151 und auf die Umfrage des Szenarios 2 134. Die demografischen Merkmale der Umfragen waren ähnlich verteilt und daher vergleichbar. Weiterhin entsprechen die Daten in etwa der Verteilung der Fans der Facebook-Seite des Mode-Einzelhandelsunternehmens (siehe Abschnitt 3).

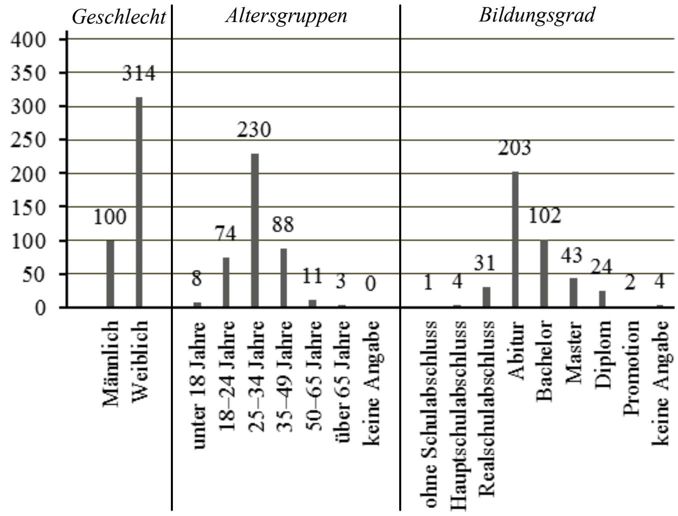


Abbildung 1. Demografische Ergebnisse in absoluten Zahlen (n=414)

In Tabelle 3 werden jeweils die auf zwei Stellen hinter dem Komma gerundeten Mittelwerte der Ergebnisse der drei Umfragen („echte“ Umfrage, Szenario 1 des experimentellen Teils der Studie sowie Szenario 2 des experimentellen Teils der Studie) dargestellt. Bei den vier Konstrukten *Informationsgehalt*, *Unterhaltung*, *Verärgerung* und *Wert des Beitrags* werden die aggregierten Mittelwerte über die drei jeweiligen Items pro Konstrukt aufgezeigt. Die Skalen haben ein Minimum von 1 und ein Maximum von 5. Um mögliche Unterschiede zwischen Szenario 1 und 2 statistisch zu identifizieren, wurde mithilfe der Statistik-Software IBM SPSS Statistics eine multivariate Varianzanalyse (MANOVA) durchgeführt. Als Ergebnis dieser Analyse werden in der Tabelle 3 die entsprechenden Signifikanzlevel der Unterschiede (p-Wert) aufgeführt.

Tabelle 3. Ergebnisse der Beiträge

Bereich		*Echt*		*Szenario 1*	*Szenario 2*	*p-Wert*
Beitrag 1 (E-Commerce; Bild; überregionale Zielgruppe)						
Wahrnehmung	Informationsgehalt	4,05		3,50	3,35	0,234 n.s.
	Unterhaltung	2,80		2,27	2,18	0,358 n.s.
	Verärgerung	1,65		2,36	2,33	0,765 n.s.
	Wert des Beitrags	3,47		2,76	2,71	0,664 n.s.
Einstellung	Einstellung zu dem Beitrag	3,70		3,26	3,19	0,642 n.s.
Reaktion	Intention: „Gefällt mir”	2,78		1,96	1,97	0,950 n.s.
	Intention: „Kommentar”	1,72		1,21	1,21	0,961 n.s.
	Intention: „Teilen”	1,74		1,46	1,51	0,625 n.s.
	Intention: „Klick auf Link”	4,00		3,42	3,38	0,786 n.s.

Bereich		*Echt*		*Szenario 1*	*Szenario 2*	*p-Wert*
Beitrag 2 (E-Commerce; Link; überregionale Zielgruppe)						
Wahrnehmung	Informationsgehalt	3,54		2,55	2,63	0,807 n.s.
	Unterhaltung	3,44		2,38	2,60	0,376 n.s.
	Verärgerung	1,62		2,64	2,62	0,710 n.s.
	Wert des Beitrags	3,36		2,36	2,55	0,791 n.s.
Einstellung	Einstellung zu dem Beitrag	3,91		2,64	2,98	0,163 n.s.
Reaktion	Intention: „Gefällt mir”	2,92		1,74	1,90	0,293 n.s.
	Intention: „Kommentar”	1,75		1,17	1,23	0,689 n.s.
	Intention: „Teilen”	1,64		1,30	1,38	0,973 n.s.
	Intention: „Klick auf Link”	3,59		2,56	2,77	0,117 n.s.
Beitrag 3 (Trends & Styling-Tipps; Bild; regionale & überregionale Zielgruppe)						
Wahrnehmung	Informationsgehalt	3,62		2,54	2,67	0,503 n.s.
	Unterhaltung	3,57		2,47	2,55	0,885 n.s.
	Verärgerung	1,58		2,60	2,47	0,340 n.s.
	Wert des Beitrags	3,50		2,35	2,45	0,748 n.s.
Einstellung	Einstellung zu dem Beitrag	3,94		2,58	2,78	0,414 n.s.
Reaktion	Intention: „Gefällt mir”	3,15		1,72	1,78	0,845 n.s.
	Intention: „Kommentar”	1,89		1,25	1,20	0,302 n.s.
	Intention: „Teilen”	1,82		1,25	1,32	0,907 n.s.
	Intention: „Klick auf Link”	3,59		2,33	2,38	0,856 n.s.
Beitrag 4 (stationäre Aktion; Bild; regionale Zielgruppe)						
Wahrnehmung	Informationsgehalt	3,88		3,66	3,23	0,001***
	Unterhaltung	3,52		3,32	3,15	0,150 n.s.
	Verärgerung	1,70		1,94	2,09	0,165 n.s.
	Wert des Beitrags	3,42		3,06	2,82	0,042*
Einstellung	Einstellung zu dem Beitrag	3,81		3,51	3,27	0,085 n.s.
Reaktion	Intention: „Gefällt mir”	3,06		2,64	2,49	0,395 n.s.
	Intention: „Kommentar”	1,83		1,52	1,45	0,540 n.s.
	Intention: „Teilen”	1,94		1,79	1,71	0,580 n.s.
Beitrag 5 (stationäre Aktion; Veranstaltung; regionale Zielgruppe)						
Wahrnehmung	Informationsgehalt	4,20		3,70	3,00	0,000***
	Unterhaltung	3,29		2,66	2,26	0,000***
	Verärgerung	1,50		2,04	2,40	0,006**
	Wert des Beitrags	3,68		3,02	2,42	0,000***
Einstellung	Einstellung zu dem Beitrag	4,11		3,41	2,69	0,000***
Reaktion	Intention: „Gefällt mir”	3,38		2,32	1,75	0,001***
	Intention: „Kommentar”	1,88		1,35	1,26	0,283 n.s.
	Intention: „Teilen”	1,89		1,62	1,34	0,016*
	Intention: „Klick auf Interessiert”	3,76		2,91	1,70	0,000***
Beitrag 6 (Entertainment; Bild; regionale & überregionale Zielgruppe)						
Wahrnehmung	Informationsgehalt	1,96		1,53	1,56	0,719 n.s.
	Unterhaltung	3,10		2,40	2,63	0,079 n.s.
	Verärgerung	1,99		2,90	2,74	0,291 n.s.
	Wert des Beitrags	2,42		1,87	1,88	0,925 n.s.
Einstellung	Einstellung zu dem Beitrag	3,17		2,31	2,51	0,185 n.s.
Reaktion	Intention: „Gefällt mir”	2,70		1,70	2,04	0,025*
	Intention: „Kommentar”	1,64		1,23	1,23	0,995 n.s.
	Intention: „Teilen”	1,74		1,32	1,36	0,681 n.s.

nicht signifikant n.s.; signifikant *; stark signifikant **; hochsignifikant***

6 Diskussion

Die Analyse zeigt, dass die Ergebnisse der „echten“ Umfrage – teils deutlich – von den Ergebnissen der Szenario-Umfragen abweichen. Diese Unterschiede zeigen sich konsistent in allen Bereichen. Die z.T. deutlichen (positiven) Abweichungen könnten auf eine starke (positive) Beeinflussung der Ergebnisse durch die Marke erklärt werden. Hieraus lässt sich folgende Hypothese ableiten:

H1: Marken beeinflussen die Wahrnehmung, Einstellung und Reaktion hinsichtlich Facebook-Beiträgen stark.

Des Weiteren zeigt die Analyse der Umfrageergebnisse, dass die Beiträge mit regionalem Bezug bei den Teilnehmenden des überregionalen Szenarios 2 schlechter bewertet werden. So sind vor allem bei Beitrag #5 – der eine reine stationäre Verkaufsveranstaltung bewerben soll – signifikante Unterschiede gemessen worden. Es kann daher die folgende Hypothese formuliert werden:

H2: Stationäre Aktionen werden von einer überregionalen Zielgruppe schlechter bewertet.

Anders war dies bei überregional ausgerichteten Beiträgen. Dort wurde die Aussage widerlegt, dass überregionale Beiträge mit E-Commerce-Inhalten von einer regionalen Zielgruppe schlechter bewertet werden. Bei beiden E-Commerce-Beiträgen #1 und #2 konnten keine signifikanten Unterschiede zwischen den Zielgruppen gemessen werden. Es folgt daraus die Hypothese:

H3: Überregionale Beiträge mit E-Commerce-Inhalten werden von einer regionalen Zielgruppe nicht schlechter bewertet.

Bei den neutralen Beiträgen #3 und #6 konnten – wie erwartet – keine signifikanten Unterschiede zwischen den Zielgruppen gemessen werden. Daher konnte folgende Hypothese abgeleitet werden:

H4: Neutrale Inhalte, wie z.B. Trends und Styling-Tipps sowie Entertainment, werden von beiden Zielgruppen ähnlich bewertet.

Aus diesen Erkenntnissen können Unternehmen mit einem ähnlichen Zielgruppendilemma schließen, dass bei überregionalen Beiträgen keine Differenzierung vorgenommen werden muss, regionale Beiträge jedoch gezielt eingesetzt werden sollten. Facebook bietet dazu bei den kostenpflichtigen Werbemöglichkeiten an, eine Zielgruppendefinition nach Standort vorzugeben. Damit sind diese Werbeanzeigen nur für Nutzer sichtbar, die sich an den vorher definierten Standorten aufhalten. Als Auswahlmöglichkeiten ist es möglich, ein oder mehrere Länder, eine Region, eine Provinz, einen Ort, einen Wahlbezirk oder eine Postleitzahl einzugeben [23]. Diese Möglichkeit beschränkt sich jedoch auf die Bewerbung eines Beitrags, jedoch nicht auf die (reguläre) Darstellung der Beiträge selbst. Daher sollte bei regionalen Beiträgen grundsätzlich überlegt werden, ob die Inhalte bzw. die Aktion „wichtig“ genug sind, um diese der ganzen Fangemeinde anzuzeigen. Folglich sollte bei diesen Beiträgen auch besonders auf den Inhalt geachtet werden.

7 Fazit und Ausblick

Ziel dieses Beitrags war es, das Zielgruppendilemma zwischen stationärem Handel und E-Commerce hinsichtlich der Marketingaktivitäten von Unternehmen auf Facebook zu untersuchen. Es wurde zunächst ein Mode-Einzelhandelsunternehmen als Fallstudie vorgestellt. Basierend auf einer Analyse des Facebook-Auftritts des Unternehmens wurde eine experimentelle Studie entworfen, um festzustellen, ob Beiträge von Kunden des Online-Shops und des (stationären) Kaufhauses unterschiedlich bewertet werden. Hierzu wurde neben demografischen Daten, das Involvement auf den Ebenen *Wahrnehmung*, *Einstellung* und *Reaktion,* bei verschiedenen Beiträgen, abgefragt. Als Ergebnis konnten vier Hypothesen abgeleitet werden.

Bei der Durchführung dieser Studie wurden erprobte und wissenschaftlich akzeptierte Methoden eingesetzt, jedoch kann die Wirkung von Werbung nur bis zu einem gewissen Grad durch herkömmliche Konsumentenbefragungen ermittelt werden. „Die durch Markenkommunikation entstehende Bedeutung der Marken ist den Kunden nicht explizit bewusst“ [24]. Zudem besteht hinsichtlich des Aufbaus des experimentellen Teils der Studie die Möglichkeit, dass die Einführung der Szenarien nicht ausreichend war, um die Teilnehmenden wie gewollt zu beeinflussen. Die Ergebnisse könnten zudem von situativen Faktoren beeinflusst worden sein, da es aufgrund des online erhobenen Fragebogens keine standardisierte Umgebung gab.

Die Ergebnisse der Umfrage könnten als Ausgangspunkt für weiterführende Studien verwendet werden. So sollten die aufgestellten Hypothesen einer konfirmativen Studie unterzogen werden. Zusätzlich könnten durch den Einsatz von Strukturgleichungsmodellen auch Korrelationen und Abhängigkeiten zwischen den Bereichen *Wahrnehmung*, *Einstellung* und *Reaktion* analysiert werden. Schließlich könnte auch der in dieser Studie deutlich erkennbare Einfluss einer Marke auf das Involvement im Social-Media-Umfeld weiter analysiert werden.

Abschließend können die Ergebnisse von Unternehmen verwendet werden, um ähnliche Zielgruppendilemmata zu adressieren und die eigenen Aktivitäten auf Social-Media-Plattformen strategisch und fundiert zu planen. Zudem kann der Ansatz in dieser Studie als Blaupause für weitere experimentelle Studien verwendet werden, um die Wahrnehmung von Social-Media-Inhalten beim Rezipienten zu untersuchen.

Referenzen

1. Heidemann, J., Klier, M., Probst, F.: Online social networks. A survey of a global phenomenon. Computer Networks 56, 3866–3878 (2012)
2. Kaplan, A.M., Haenlein, M.: Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 53, 59–68 (2010)
3. Kietzmann, J.H., Hermkens, K., McCarthy, I.P., Silvestre, B.S.: Social media? Get serious! Understanding the functional building blocks of social media. Business Horizons 54, 241–251 (2011)
4. Wilkinson, Z.: Oh, how Pinteresting! An introduction to Pinterest. Library Hi Tech News 30, 1–4 (2013)

5. Boyd, D.M., Ellison, N.B.: Social Network Sites. Definition, History, and Scholarship. Journal of Computer-Mediated Communication 13, 210–230 (2007)
6. Smith, A.N., Fischer, E., Yongjian, C.: How Does Brand-related User-generated Content Differ across YouTube, Facebook, and Twitter? Journal of Interactive Marketing 26, 102–113 (2012)
7. Bruhn, M., Schäfer, D.B., Schwarz, J., Lauber, M.: Facebook, Twitter, YouTube und Co. – Erwartungen der Nutzer an Social-Media-Plattformen. Marketing Review St. Gallen 28, 36–42 (2011)
8. Bruhn, M., Martin, S., Schnebelen, S.: Integrierte Kommunikation in der Praxis. Entwicklungsstand in deutschsprachigen Unternehmen. Springer Gabler, Wiesbaden (2014)
9. Handelsverband Deutschland (HDE): E-Commerce-Umsätze, http://www.einzelhandel.de/index.php/presse/zahlenfaktengrafiken/item/110185-e-commerce-umsaetze (zuletzt zugegriffen am: 02.12.2016)
10. Handelsverband Deutschland (HDE): Onlineshops des stationären Handels, http://www.einzelhandel.de/index.php/presse/zahlenfaktengrafiken/item/122998-onlineshops-des-station%C3%A4ren-handels (zuletzt zugegriffen am: 10.07.2016)
11. Verma, R., Jahn, B., Kunz, W.: How to transform consumers into fans of your brand. Journal of Service Management 23, 344–361 (2012)
12. Lorenzo-Romero, C., Constantinides, E., Alarcón-del-Amo, M.: Consumer adoption of social networking sites. Implications for theory and practice. Journal of Research in Interactive Marketing 5, 170–188 (2011)
13. Vries, L. de, Gensler, S., Leeflang, P.S.: Popularity of Brand Posts on Brand Fan Pages. An Investigation of the Effects of Social Media Marketing. Journal of Interactive Marketing 26, 83–91 (2012)
14. Bolotaeva, V., Cata, T.: Marketing Opportunities with Social Networks. JISNVC, 1–8 (2011)
15. Meffert, H., Burmann, C., Kirchgeorg, M.: Marketing. Grundlagen marktorientierter Unternehmensführung Konzepte - Instrumente - Praxisbeispiele. Springer Gabler, Wiesbaden (2015)
16. Castronovo, C., Huang, L.: Social Media in an Alternative Marketing Communication Mode. Journal of Marketing Development & Competitiveness 6, 117–136 (2012)
17. Heller-Baird, C., Parasnis, G.: From social media to social customer relationship management. Strategy & Leadership 39, 30–37 (2011)
18. Trommsdorff, V.: Konsumentenverhalten. Kohlhammer, Stuttgart (2009)
19. Ducoffe, R.H.: Advertising value and advertising on the web. JOURNAL OF ADVERTISING RESEARCH 36, 21–35 (1996)
20. Alhabash, S., McAlister, A.R., Lou, C., Hagerstrom, A.: From Clicks to Behaviors. The Mediating Effect of Intentions to Like, Share, and Comment on the Relationship Between Message Evaluations and Offline Behavioral Intentions. Journal of Interactive Advertising 15, 82–96 (2015)
21. Brackett, L.K., Carr, B.N., Jr.: Cyberspace advertising vs. other media: Consumer vs. mature student attitudes. JOURNAL OF ADVERTISING RESEARCH 41, 23–32 (2001)
22. Burton, S., Lichtenstein, D.R.: The Effect of Ad Claims and Ad Context on Attitude Toward the Advertisement. Journal of Advertising 17, 3–11 (1988)
23. Facebook: Hilfebereich, https://www.facebook.com/business/help/633474486707199/ (zuletzt zugegriffen am: 02.12.2016)
24. Held, D., Scheier, C.: Wie Werbung wirkt. Erkenntnisse des Neuromarketing. Haufe Verlag, München (2013)

Co-Creation and User Involvement in a Living Lab: An Evaluation of Applied Methods

Theodor Beutel, Julia M. Jonas, Kathrin M. Moeslein

Friedrich-Alexander-University Erlangen-Nuremberg, Information Systems 1, Nuremberg, Germany
{theodor.beutel;julia.jonas;kathrin.moeslein}@fau.de

Abstract. Living labs are only recently developing to facilitate active user involvement in an interactive setting. Research on the methodological facilitation of co-creation and user feedback in such open physical spaces is still scarce. The objectives of this paper are to identify applied methods as well as to investigate the level of user involvement in living labs to further develop theoretical insights on living labs as well as on method implementations for co-creation. A qualitative explorative approach in the form of a case study on the living lab JOSEPHS in Nuremberg is applied. This paper finds that applied methods serve either of two purposes: 1) Collecting data for innovation research, or 2) adapting co-creation to living labs. Combined accordingly, methods cover both purposes and increase user involvement. Furthermore, six factors that determine user involvement are proposed. Implications for living lab managers are provided.

Keywords: Co-Creation, Living Labs, User Involvement, Methods, ICT

1 Introduction

In an era of sophisticated information and communications technology (ICT) with empowered users and blurring organisational boundaries, innovation procedures in new product and service development (NPSD) undergo a fundamental transformation. Firms actively loosen conventional boundaries through the inclusion of external actors in their NPSD activities [1]. As part of such open innovation approaches, major importance is devoted to users. Whereas users only recently received major attention in open innovation research [2], the emergence of user innovation research dates back as far as four decades [3, 4]. Among open innovation practices, co-creation with users is one of the most important and proves to be widely adopted among firms [5].

Today, many tools of open innovation are driven by ICT [6]. Benefiting from low cost and large scale, also co-creation is often implemented online [7]. However, building trust [7] and providing context [8] proves to be challenging over the Internet. Here, real life settings play to their strengths [9, 10]. Living labs (LL), endorsed by the

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Beutel, T.; Jonas, J.M.; Moeslein, K.M. (2017): Co-Creation and User Involvement in a Living Lab: An Evaluation of Applied Methods, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1453-1464

European Union [11], are a rather new approach to foster NPSD and co-creation with users in real life settings. LLs are described as a "new way to manage the new product development process" [12] and fit into the idea of arenas for co-creation by Bhalla [10]. LL are considered as a conceptualisation of extra-organisational open innovation [2, 8].

However, it remains rather unclear *how* users can and should be involved in the context of LLs. In contrast to the firms' and the users' roles, the role of intermediaries in open innovation is less illuminated [13, 14]. Furthermore, recent innovation research has covered broadly *where* but less *how* to search for external knowledge [15]. It is particularly questioned how methods involving the user in a LL should be chosen and implemented. Such lack of knowledge may be attributed to the novel nature of LLs as an area of interest. However, the importance of user involvement specifically in LLs has been highlighted by scholars for more than a decade [16]. Mulder and Stappers [17] state that LLs are not living up to their full potential of active user involvement in real life settings. Hence, this paper aims to identify applied methods and to investigate the emergence and intensity.

2 Theoretical Foundation

2.1 Co-Creation for Innovation Purposes

As real and physical spaces, LLs can be used for open innovation by facilitating co-creation with users [2, 18, 19]. Co-creation for innovation purposes in NPSD is only one out of several applications for co-creation [20] and is framed within the concept of co-creation of value. While traditionally, the process of value creation was coined by independently value-adding firms, which led to demand from passively consuming users, value is now jointly co-created from firms and users. When users go beyond mere consumption and become active contributors in NPSD, they co-create and extract value for their own good [21]. Through the active role of users, the changed user-firm relationship implies a new locus of value creation, which lies in interaction and experience [22].

In this paper, co-creation is defined as "an active, creative and social collaboration process between producers[1] (retailers) and customers[2] (users)" [24]. It is argued that users place importance primarily on the value that (eventually) emerges while the process is of little account to them [29]. Consequently, Snyder et al. [29] propose to view outcome and process separately. However, it is not to be neglected that users create value during co-creation processes [30], e.g. benefit from enjoyment and learning [26]. It is therefore useful to shed more light on processes and methods of co-creation.

For a firm, the most important benefit of co-creation with users lies in an improved access to need information, as need information tends to be sticky. Thus, better access to user preferences leads to more effective NPSD [24].

[1] The literature contains varying terminology, such as "company" [as in 21, 23] and "producer" [as in 3, 24]. This paper confines the word choice to the term "firm" [as in 1, 2].

[2] Likewise with footnote 1, this paper employs the term "user" [as in 3, 25] in place of "consumer" [as in 26, 22], "customer" [as in 21, 24], "citizen" [as in 27, 28] and "visitor" [as in 23].

2.2 User Involvement

The involvement[3] of users is a fundamental dimension to co-creation [22] as well as to open innovation [2]. User involvement can be defined as a user's influence on the idea, development and launch processes in NPSD [25]. Similarly, Piller, Ihl, and Vossen [24] see co-creating users actively involved during NPSD processes, but add that they are performing "an act of company-to-customer interaction which is facilitated by the company" [24]. Depending on their role, users engage in different intensity, varying in time and effort. User activity in NPSD processes can range widely among users as a passive source of information, a co-creating contributor, and a designing innovator [33].

2.3 Living Labs

LLs provide a novel way to connect firms with users and help with "closing the gap between open and user innovation" [34]. Compared to other innovation approaches, LLs differ in two dimensions [8]. Firstly, LLs are capable of providing novel structures for user involvement [35]. They involve users in an interactive and empowering way, enabling them to become co-creators, and thus go beyond user-centred approaches that only involve users passively [36]. Secondly, a particular strength of LLs lies in their real life offline setting. In this regard, they overcome hurdles in knowledge transmission relating to sticky information and tacit knowledge [37]. Therefore, LLs can be considered as "a user-centric research methodology for sensing, prototyping, validating and refining complex solutions in multiple and evolving real life contexts" [38].

Methods applied in Living Labs. Bergvall-Kåreborn and Ståhlbröst [27] argue that LLs are methodologically designed in two levels. While a general framework provides stability and continuity, a second level beneath allows spontaneity within projects [27]. Similarly, Dell'Era and Landoni [12] argue that a framework both allows and determines the implementation of methods within a LL. The LL methodology draws from co-creation techniques as well as from traditional innovation research methods such as questionnaires, in-depth interviews, or focus groups. Depending on the domain and its method, a different type of knowledge and originality can be expected [39].

Distinguishing itself from other approaches, the LL methodology stands out in active user involvement and realism [17, 37]. Real life environments set LLs apart from controlled environments. A real life setting is usually designed through the use of contextual methods and/or with the aid of physical artefacts [12]. It is argued that methods applied in LLs should adapt to the distinct advantages in interactivity and real life environments and thus should go beyond traditional methods of innovation research [35, 40]. However, only few studies evaluate methods applied in LLs in relation to the distinct features of LLs, whereas traditional methods have been researched extensively [27]. All in all, distinct attributes of LLs and advantages over other innovation approaches have been emphasised sufficiently [8, 9, 35, 37]. However, studies linking

[3] This paper uses the term "involvement" [as in 25, 30, 31], whereas the literature interchangeably employs the terms "participation" [as in 3, 26] and "integration" [as in 24]. However, it is not referred to committing oneself emotionally as in "commitment" or "dedication" [32].

methods applied to LLs to these attributes are sparse and tend to focus on NPSD phases [18, 41] or single methods [42, 43]. Consequently, this paper conducts a comprehensive analysis of methods and their characteristics. Accordingly, it is questioned:

(1) *Which methods are implemented in a living lab and how are they characterised according to the level of user involvement?*

User Involvement in a Living Lab. While involving users is only one factor among many that promote co-creation in a LL [19, 44], it is considered quintessential to the LL concept [12, 35]. In LLs, firms are often one stakeholder among many [45]. Hence distinguishing between stakeholders is required. Due to LL's roles as intermediaries in innovation, Piller, Ihl, and Vossen's [24] understanding of a *firm-user interaction* is altered to *intermediary-user interaction*. This adjustment seems appropriate because the co-creation process takes place *in* and *with* the LL, acting as an agent for the firm. Nevertheless, the firm is still considered to facilitate the co-creation process [24].

Regarding voluntary user involvement, many questions remain for future research [46]. A dearth of methods and tools adapted to the distinct attributes of LLs has been emphasised [27]. As a first step, methods applied in a LL are examined for the level of user involvement [28]. As a result, it is proposed to shift from user-centred to user-driven methods. However, Gray et al. [28] do not present further implications concerning the application of co-creation methods in LLs. To date, only few studies evaluate methods applied in LLs specifically in relation to user involvement as a distinct feature of LLs, while observed LLs cases vary widely [28, cf. 40, 41, 47]. Hence, the second research question of this paper is as follows:

(2) How does user involvement differ and how is it determined?

3 Research Design

A qualitative explorative approach in the form of a holistic single-case study is applied. To gain in-depth insights on co-creation methods and user involvement in LLs, qualitative case study research is a suitable methodology [48]. Further reason lies in the opportunity to illuminate contextual conditions and processes [48, 49]. The LL serving as unit of analysis of the present case providing a unique setting and the LL landscape being rather diverse [12, 50] further justifies a single-case design [48]. Hence, this paper focusses on an in-depth analysis rather than aspiring general claims [48].

3.1 The Case

The case study is implemented at JOSEPHS® – Die Service Manufaktur, a LL centrally situated in Nuremberg in southern Germany. Within its premises, which also include a coffee shop and a workshop area, the LL devotes an openly accessible area to five distinct co-creation spaces, used by five companies simultaneously. Six days a week during regular shopping hours, any passer-by is invited to come in; LL visitors await the opportunity to engage themselves interactively in firms' NPSD processes. Just as

LL visitors are expected to share their feedback, companies are advised to be equally open and cooperative. Firms can utilise JOSEPH's real life environment to test (physical and digital) ideas and prototypes under simulated conditions with a diverse, self-selected crowd of users. Since the LL's launch in 2014, users had been able to co-create about 60 diverse products and services at the LL. The firms utilising the LL come from a broad variety of backgrounds, ranging from start-ups in consumer products, to technology providers and larger enterprises even in business-to-business industries.

3.2 Data Collection and Analysis

This case study took place during summer 2016 and is based on primary data in the form of physical artefacts and seven semi-structured expert interviews as well as on secondary data from documentation material such as reports and photo documentation.

To utilise the expert's knowledge effectively and ensure comparability, an interview guide allowed open responses within a predefined field of interest [48, 51, 52]. All interviews were audio recorded, transcribed following the rules proposed by Misoch [52], and analysed using the qualitative data analysis software MAXQDA [49, 53]. The challenges of quality in interview data lie in potential biases, poor recall, or inaccurate articulation [48]. These are addressed through a diverse sampling of interview partners who occupy various perspectives and positions in three different organisations as well as through complementary data from documentation and artefacts [51].

While the interviews constitute the main and most important source of information, including documentation and artefacts both forms a reliable starting point for the case and allows to verify and contextualise interview data in a complementary way [48, 49]. Press releases, photo documentation and various publicity materials as well as internal documents by the LL are analysed. Due to the importance of context, physical artefacts such as the LL itself and objects within the LL are included in this case [54].

As part of the data analysis process, raw data from all three data sources was approached through open coding and iteratively complemented with existing literature [49, 51, 55]. Upon completion of the coding process, all codes, code segments and comments were exported from MAXQDA to Microsoft Excel for further analysis.

4 Empirical Findings

A range of applied methods is identified. The variety stems from a discrepancy in purpose, as shown in Table 1. Whereas one group of methods is utilised primarily to have the user answer specific questions, the other group of methods primarily aims to stimulate the user's experience in the LL. Ultimately, all methods serve the purpose of innovation research. While the former contribute directly to data collection, the latter do indirectly – complementing the former with beneficial LL-specific characteristics.

Methods of innovation research are threefold. Questioning methods of quantitative and qualitative nature as well as observational methods including technology-assisted tracking are applied. Some of these traditional innovation research methods such as questionnaires or voting mechanisms involve the user in a rather passive and theoretical way. In order to involve the user in a more active and practical way, a single co-creation space can be equipped with several complementary methods, as one expert phrased as follows: "*[If I] should test something, or tinker with something, then there is an active involvement which is what we want, but which can be achieved through [traditional] innovation research methods only then, if there is an app to try out, for instance.*"

Purpose	Innovation Research			Complementary
Type	Quantitative questioning	Qualitative questioning	Observing	
Methods	Price assessments, Questionnaires, Usability-tests (quantitative), Voting	Focus Groups, Interviews, Open feedback, Personas, Usability-tests (qualitative), Workshops	Observation and shadowing, Tracking	Artefacts, Storytelling, Information material, Toolkits, Prototypes, Prototyping, Service staging

Table 1: Characterisation of Applied Methods

The experts consider involving users actively as crucial, one stating that "*we try to involve the user as much as possible, so you would rarely see a yes-no-query as the only method, but rather as a supplement.*" It is also described as a prerequisite to generating data for innovation research. Furthermore, it does not only make it easier for users to give feedback but also increases their willingness to do so. Methods differ regarding the facilitation of active user involvement in methods, as Figure 1 illustrates.

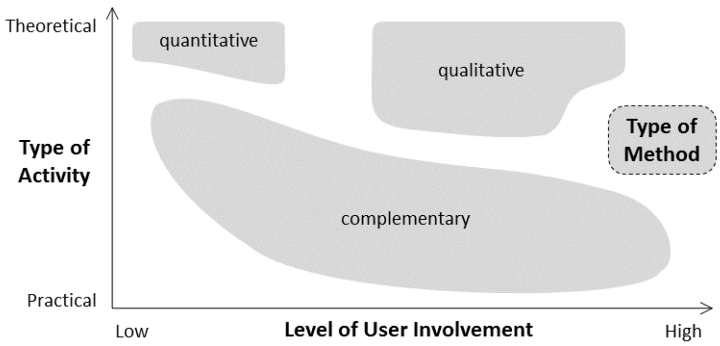


Figure 1: Differing User Involvement and Activity in Applied Methods

If considered individually, quantitative and qualitative methods of innovation research involve the user predominantly on a theoretical level. They require the user to answer questions, hence to think about a certain subject. Quantitative methods result in a low to medium level of user involvement in terms of activity and variety. Qualitative methods, on the other hand, rank substantially higher in user involvement. However, they are equally constrained to theoretical activities. Methods with beneficial LL characteristics may complement these innovation research methods. Examples include the provision of context and haptic experience in the case of physical artefacts, and testable prototypes with a high level of practical activity, as two experts explain: "*Whenever possible, we hand something over to the user […] if it is a physical thing or so*", because "*through mere haptic experiences, [the user] becomes more involved*".

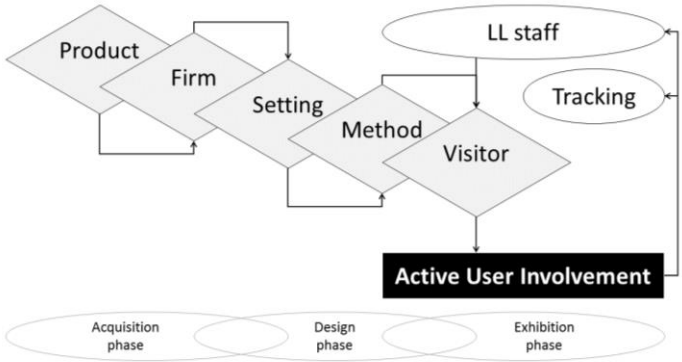


Figure 2: Factors and Conditions Determining Active User Involvement

Figure 2 outlines several factors, which are proposed to influence the emergence of active user involvement beforehand. In this model, each of the six factors product, firm, setting, method, LL staff, and users are dependent on all previous factors, with several conditions respectively as shown in Table 2.

Factor	Product	Firm	Setting	Method	User	LL staff
Conditions	Fit for co-creation, Functio-nality	Strategic fit, Willing-ness	Context/ artefacts, Appeal, No barriers	Facilitating activity	Type, Effort, Time, Motivation	Enabling/ motivating

Table 2: Factors Determining Active User Involvement

5 Discussion

Contributing to the discussion about method application and development for new methods in LLs [27, 40], this paper proposes a combinatorial approach. Several complementary methods lead to the desired outcome of data for innovation research in a more effective and more appropriate way than traditional innovation research methods on their own. Complementing methods can adapt to the distinct features of LLs and thus benefit traditional methods of innovation research through more richness and quality in data. These findings are in line with Tang and Hämäläinen [41] who argue that combining methods can lead to a better understanding of users.

A predominant use of traditional innovation research methods like interviews and surveys in LLs is observed, congruent with Tang and Hämäläinen [56]. However, it is frequently argued that ICT is underused in LLs [7, 9, 18, 41, 57, 58]. Use of ICT, however, could contribute to an improved realisation through shortening the feedback loop between users and corporate designers in iterative processes [58], enabling a true firm-user interaction beyond the intermediating LL. Equally, ICT could connect users with LLs when they are not on site. Either way, higher user involvement can be expected, as the level of activity would increase both time and effort spent.

While no systematic approach to using ICT was identified, general technology use as part of the co-creation process occurs at times. In line with previous authors, this study finds technologically sophisticated methods and tools being used infrequently besides mobile handsets for usability-tests. Counterexamples include a 'thumbs up'-voting mechanism using a Microsoft Kinect camera for gesture recognition and virtual reality (VR) applications using a Google Cardboard. Besides, tracking technology is implemented, which the user, however, is not concerned with as this is an unobtrusive way of collecting data [41]. Notably, it is found that technology use may also impede user involvement. If a lack in technology acceptance, unfamiliarity, or technical failure is present, this can keep users from becoming actively involved in the envisaged way.

This study proposes users as the chronologically last, thus deciding factor in involving themselves actively. Holding the position of a co-creating partner, LL visitors are entitled to both include and exclude themselves from the co-creation process whenever they want and without having to give reasons [59]. In an open LL, the idea of incentivising outsiders to become a LL visitor suggests itself at first sight. Dutilleul, Birrer, and Mensink [45] even ask if incentives are needed in order to attract and sustain a desired amount of LL visitors continuously. However, based on the principle of self-selection, this study argues that such mechanisms may distort results and question their validity. Previous studies find that material and financial rewards are not important to users [46] while importance is attached to the value that users experience in the course of co-creation [60]. Instead of giving financial incentives, it is proposed to improve on the co-creation space, particularly the appeal of its setting in order to provide an experience to LL visitors that is as much beneficial and pleasant as possible.

Managerial Implications

Implementing digital technology for an automated acquisition of implicit, behavioural data as well as explicit, articulated data would arguably not only support LL staff and let them focus on interpersonal communication but also accelerate the data analysis process, thus increasing efficiency. For example, a customised mobile handset that runs a digital content management system could record interviews, capture questionnaires, and aggregate contextual data (e.g. place, date, time, duration). Most importantly, LL staff needs to be provided with enough expertise in order to carry out semi-structured interviewing and other methods of qualitative innovation research.

In the design process of co-creative activities, a firm's requirements serve as the starting point and may imply certain methods. However, some firms have false expectations or request mostly quantitative methods. During communication with customer firms, it is advisable to follow a threefold strategy of selecting appropriate firms fitting the LL methodology, undertaking expectation management on the innovation research outcome, and consulting firms in order to utilise a LL's strengths.

Particularly in iterative or continuous NPSD processes in LL, the use of ICT may be of help. ICT could bridge the gap to a firm without the necessity of sending an employee physically to a LL. With an employee being available on call, all LL visitors had the chance to deepen their input. On the other hand, ICT are able to reach LL visitors before and after their physical co-creation engagement. For instance, a web interface could be a way for LL visitors to contribute even in hindsight. It does not seem too farfetched to assume that users might come up with new ideas after they left and had the opportunity to rethink their contribution, but also find their additional input not worth a second (physical) visit. In terms of a real life setting, storytelling proves beneficial for creating overall context. Combining ICT and the method of storytelling with augmented reality (AR), Snapchat, a story-based social media application, could fit LLs with changing themes quite well. Its contribution to user involvement in activity and variety is yet to be evaluated. However, it might be a way to involve younger users in particular.

6 Conclusion and Outlook

The first contribution of this paper is a review on methods that are applied in LLs. Characterising these, two primary purposes emerge. The first group of methods directly contributes to data collection for innovation research, while the second group contributes indirectly and primarily complements former methods with beneficial LL-specific characteristics in providing a real life setting and enabling user involvement.

The second contribution of this paper is a proposed model of six consecutive factors and several conditions that influence user involvement. It is proposed that these factors influence user involvement firstly in its emergence, and secondly in its intensity. If a factor allows user involvement at all, it further limits its maximum intensity that subsequent factors are able to draw on.

Concerning this study, several limitations apply. Firstly, the form of a holistic single-case study induces an in-depth analysis, which does not permit generalisation but requires comparison across multiple cases of LLs. Secondly, while this paper focuses on contributing to the intermediary perspective of LLs and is conceptualised accordingly, both the user and the firm perspective are not particularly addressed.

Exploring the field as a first step, this study indicates promising niches worth further quantitative research. Above all, the proposed model on factors influencing user involvement demands quantitative validation with proof of causal effects. Furthermore, future studies should ask whether combinations of methods with a higher level of user involvement result in superior validity or efficiency [35]. Here, studies should contribute with other perspectives than the intermediary's. As part of the firm, the rates of adoption for further development and profitability might contribute to the question which level of user involvement is considered ideal. Assuming a high level of active user involvement, it is of interest to know which kinds of firms benefit the most. Simulating and enhancing a real life setting, new technologies, such as AR and VR, seem promising. It should be evaluated if they are beneficial to the level of user involvement as well as to the co-creation process as a whole. Do these technologies lead to more motivation and willingness among LL visitors through improved experience and a higher perceived value? Although LLs differentiate themselves with their offline real life settings from Internet-based technologies, ICT usage within LLs is worth further research. While LLs utilise the Internet only rarely [35], the combination of online-offline methods could lead the way for the future of LLs. Finally, in providing *Co-Creation as a Service*, the intermediary perspective needs further research [14]. In line with Schweitzer, Gassmann, and Rau [31], it is argued that reciprocal effects of goal setting, chosen methods, and user involvement require further qualitative and quantitative modelling.

7 Acknowledgments

We thank all interview partners, the JOSEPHS team, and the anonymous reviewers for their valuable contributions to this research paper. The paper is part of current research projects funded by the Fraunhofer IIS-SCS within its "Service Innovation" initiative.

References

1. Chesbrough, H.W., Bogers, M.: Explicating Open Innovation: Clarifying an Emerging Paradigm for Understanding Innovation. In: Chesbrough, H.W., Vanhaverbeke, W., West, J. (eds.) New Frontiers in Open Innovation. OUP Oxford, Oxford (2014)
2. Bogers, M., Zobel, A.-K., Afuah, A., Almirall, E., Brunswicker, S., Dahlander, L., Frederiksen, L., Gawer, A., Gruber, M., Haefliger, S., et al.: The Open Innovation Research Landscape. Established Perspectives and Emerging Themes across Different Levels of Analysis. Industry and Innovation Forthcoming (2017)
3. von Hippel, E.: Free Innovation. MIT Press, Cambridge, MA (2017)
4. von Hippel, E.: The dominant role of users in the scientific instrument innovation process. Research Policy 5, 212–239 (1976)
5. Chesbrough, H.W., Brunswicker, S.: A Fad or a Phenomenon? Research-Technology Management 57, 16–25 (2014)
6. Möslein, K.M.: Open Innovation: Actors, Tools, and Tensions. In: Huff, A.S., Möslein, K.M., Reichwald, R. (eds.) Leading Open Innovation, pp. 69–86. MIT Press, Cambridge, MA (2013)
7. Rayna, T., Striukova, L.: Open innovation 2.0: Is co-creation the ultimate challenge? International Journal of Technology Management 69, 38–53 (2015)
8. Schuurman, D., Lievens, B., De Marez, L., Ballon, P.: Towards optimal user involvement in innovation processes: A panel-centered Living Lab-approach. Proceedings of Technology Management for Emerging Technologies, 2046–2054 (2012)
9. Tang, T., Wu, Z., Hämäläinen, M., Ji, Y.: From Web 2.0 to Living Lab. An Exploration of the Evolved Innovation Principles. Journal of Emerging Technologies in Web Intelligence 4, 379–385 (2012)
10. Bhalla, G.: How to plan and manage a project to co-create value with stakeholders. Strategy and Leadership 42, 19–25 (2014)
11. European Commission: Living Labs for User-Driven Open Innovation. An Overview of the Living Labs Methodology, Activities and Achievements. EUR-OP, Luxembourg (2008)
12. Dell'Era, C., Landoni, P.: Living Lab: A Methodology between User-Centred Design and Participatory Design. Creativity and Innovation Management 23, 137–154 (2014)
13. Howells, J.: Intermediation and the role of intermediaries in innovation. Research Policy 35, 715–728 (2006)

14. Hakkarainen, L., Hyysalo, S.: How Do We Keep the Living Laboratory Alive? Learning and Conflicts in Living Lab Collaboration. Technology Innovation Management Review 3, 16–22 (2013)
15. Lopez-Vega, H., Tell, F., Vanhaverbeke, W.: Where and how to search? Search paths in open innovation. Research Policy 45, 125–136 (2016)
16. Pierson, J., Lievens, B.: Configuring Living Labs For A 'Thick' Understanding Of Innovation. Proceedings of The Ethnographic Praxis in Industry Conference, 114–127 (2005)
17. Mulder, I., Stappers, P.J.: Co-creating in practice: Results and challenges. Proceedings of The 15th IEEE International Technology Management Conference, 1–8 (2009)
18. Feuerstein, K., Hesmer, A., Hribernik, K.A., Thoben, K.-D., Schumacher, J.: Living Labs: A New Development Strategy. In: Schumacher, J. (ed.) European living labs: A new approach for human centric regional innovation. wvb, Berlin (2008)
19. Greve, K., Martinez, V., Jonas, J.M., Neely, A., Möslein, K.M.: Facilitating co-creation in living labs: The JOSEPHS study. Cambridge Service Alliance Working Paper Series (2016)
20. Alves, H., Fernandes, C., Raposo, M.: Value co-creation. Concept and contexts of application and study. Journal of Business Research 69, 1626–1633 (2016)
21. Prahalad, C.K., Ramaswamy, V.: Co-opting Customer Competence. Harvard Business Review 78, 79–90 (2000)
22. Prahalad, C.K., Ramaswamy, V.: Co-creation experiences. The next practice in value creation. Journal of Interactive Marketing 18, 5–14 (2004)
23. Roth, A., Jonas, J.M., Fritzsche, A., Danzinger, F., Möslein, K.M.: Spaces for Value Co-Creation: The Case of "JOSEPHS® - The Service Manufactory". Proceedings of Euram 2015 (2015)
25. Jespersen, K.R.: User-Involvelement and Open Innovation: The Case of Decision-Maker Openness. International Journal of Innovation Management 14, 471–489 (2010)
26. Raasch, C., von Hippel, E.: Innovation Process Benefits: The Journey as Reward. MIT Sloan Management Review 55, 33–39 (2013)
27. Bergvall-Kåreborn, B., Ståhlbröst, A.: Living Lab: an open and citizen-centric approach for innovation. International Journal of Innovation and Regional Development 1, 356–370 (2009)
28. Gray, M., Mangyoku, M., Serra, A., Sánchez, L., Aragall, F.: Integrating Design for All in Living Labs. Technology Innovation Management Review 4, 50–58 (2014)
29. Snyder, H., Witell, L., Gustafsson, A., Fombelle, P., Kristensson, P.: Identifying categories of service innovation. A review and synthesis of the literature. Journal of Business Research 69, 2401–2408 (2016)
30. Nambisan, S., Baron, R.A.: Interactions in virtual customer environments. Implications for product support and customer relationship management. Journal of Interactive Marketing 21, 42–62 (2007)
31. Schweitzer, F., Gassmann, O., Rau, C.: Lessons from Ideation: Where Does User Involvement Lead Us? Creativity and Innovation Management 23, 155–167 (2014)
32. Merriam-Webster: Definition of Involvement, http://www.merriam-webster.com/dictionary/involvement.
33. Cui, A.S., Wu, F.: Utilizing customer knowledge in innovation. Antecedents and impact of customer involvement on new product performance. Journal of the Academy of Marketing Science 44, 516–538 (2016)
34. Schuurman, D.: Bridging the gap between Open and User Innovation? Ghent (2015)
35. Almirall, E., Wareham, J.: Living Labs and open innovation: roles and applicability. The Electronic Journal for Virtual Organizations and Networks 10, 21–46 (2008)
36. Lehmann, V., Frangioni, M., Dubé, P.: Living Lab as knowledge system: An actual approach for managing urban service projects? Journal of Knowledge Management 19, 1087–1107 (2015)

37. Almirall, E., Lee, M., Wareham, J.: Mapping Living Labs in the Landscape of Innovation Methodologies. Technology Innovation Management Review 2, 12–18 (2012)
38. Eriksson, M., Niitamo, V.-P., Kulkki, S.: State-of-the-Art in utilizing Living Labs approach to user-centric ICT innovation – a European approach. Centre for Distance-Spanning Technology at Luleå University of Technology, Sweden (2005)
39. Witell, L., Kristensson, P., Gustafsson, A., Löfgren, M.: Idea generation. Customer co-creation versus traditional market research techniques. Journal of Service Management 22, 140–159 (2011)
40. Franz, Y.: Designing social living labs in urban research. info 17, 53–66 (2015)
41. Tang, T., Hämäläinen, M.: Beyond Open Innovation: the Living Lab Way of ICT Innovation. Interdisciplinary Studies Journal 3, 15–23 (2014)
42. Logghe, S., Schuurman, D.: Action Research as a framework to evaluate the operation of a Living Lab. Proceedings of OpenLivingLab Days 2016, 186–194 (2016)
43. Coenen, Tanguy: Towards FALL: a Framework for Agile Living Lab projects. Proceedings of OpenLivingLab Days 2016, 129–144 (2016)
44. Mulder, I., Velthausz, D., Kriens, M.: The Living Labs Harmonization Cube: Communicating Living Labs' Essentials. The Electronic Journal for Virtual Organizations and Networks 10, 1–14 (2008)
45. Dutilleul, B., Birrer, F.A.J., Mensink, W.: Unpacking European Living Labs: Analysing Innovation's Social Dimension. Central European Journal of Public Policy 4, 60–85 (2010)
46. Logghe, S., Baccarne, B., Schuurman, D.: An exploration of user motivations for participation in Living Labs. Proceedings of The XXV ISPIM Conference – Innovation for Sustainable Economy & Society, 1–10 (2014)
47. Verilhac, I., Pallot, M., Aragall, F.: IDeALL: Exploring the way to integrate design for all within living labs. Proceedings of The 18th International Conference on Engineering, Technology and Innovation 2012, 1–8 (2012)
48. Yin, R.K.: Case study research. Design and methods. SAGE, Los Angeles (2009)
49. Flick, U.: An Introduction to Qualitative Research. SAGE, Los Angeles (2009)
50. Leminen, S.: Living Labs as Open Innovation Networks. Networks, Roles and Innovation Outcomes. Aalto University, Aalto (2015)
51. Bogner, A., Littig, B., Menz, W.: Interviews mit Experten. Eine praxisorientierte Einführung. Springer, Wiesbaden (2014)
52. Misoch, S.: Qualitative Interviews. De Gruyter, Berlin (2015)
53. Miles, M.B., Huberman, A.M., Saldana, J.: Qualitative data analysis: A methods sourcebook. SAGE (2013)
54. Gillham, B.: Case study research methods. Continuum, London (2000)
55. Taylor, S.J., Bogdan, R., DeVault, M.L.: Introduction to qualitative research methods. A guidebook and resource. Wiley, Hoboken, NJ (2016)
56. Tang, T., Hämäläinen, M.: Living Lab Methods and Tools for Fostering Everyday Life Innovation. Proceedings of The 18th International Conference on Engineering, Technology and Innovation 2012, 1–8 (2012)
57. Ståhlbröst, A., Sällström, A., Holst, M.: User Evaluations in the Wild – Experiences from Mobile Living Labs. Proceedings of Mobile Living Labs 09, 1–4 (2009)
58. Følstad, A.: Towards a Living Lab for the Development of Online Community Services. The Electronic Journal for Virtual Organizations and Networks 10, 47–58 (2008)
59. Ståhlbröst, A.: Forming future IT : the living lab way of user involvement. Luleå (2008)
60. Georges, A., Schuurman, D., Vervoort, K.: Factors Affecting the Attrition of Test Users During Living Lab Field Trials. Technology Innovation Management Review 6, 35–44 (2016)

Advancing the Adoption of a New Generation of Certifications – A Theoretical Model to Explain the Adoption of Continuous Cloud Service Certification by Certification Authorities

Andrea Quinting[1], Sebastian Lins[1], Jakub Szefer[2], and Ali Sunyaev[3]

[1] University of Cologne, Department of Information Systems, Cologne, Germany
{quinting,lins}@wiso.uni-koeln.de

[2] Yale University, Department of Electrical Engineering, New Haven, CT, U.S.
jakub.szefer@yale.edu

[3] University of Kassel, Research Center for IS Design (ITeG), Kassel, Germany
sunyaev@uni-kassel.de

Abstract. Cloud certifications are a good means to assure users of high level of security and reliability of certified cloud services. However, cloud environments are highly dynamic due to the challenging cloud characteristics and fast technology life-cycles. We believe that current certifications fail to cope with an ever-changing cloud environment because assessments are based only on manual expert assessments and periodic spot checks. We argue that continuous service certification (CSC) is required to assure reliable and trustworthy cloud services. To understand and enhance CSC's rate of adoption, we examine the adoption process of CSC from the perspective of certification authorities by building on the Diffusion of Innovations theory and the Technology-Organization-Environment framework. Our findings reveal that the innovation's characteristics, organizational and environmental influences will affect the adoption of CSC by certification authorities. We advance the understanding of the CSC adoption process by providing a synthesis and discussion of important factors.

Keywords: Continuous Certification, Cloud Services, Diffusion of Innovations Theory, Technology-Organization-Environment Framework

1 Introduction

Several cloud service certifications have recently evolved and attempt to assure users of a high level of security, availability and legal compliance of the certified cloud service. Certifications aim to reduce cloud customers' concerns, increase trust as well as to enhance transparency in the cloud service markets. These cloud service markets have become increasingly popular because they offer a vast selection of IT services (e.g., online storage, office software and collaboration tools) that are instantly available and that can withstand unexpected fluctuations in demand for the service, e.g., quickly spin-

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Quinting, A.; Lins, S.; Szefer, J.; Sunyaev, A. (2017): Advancing the Adoption of a New Generation of Certifications – A Theoretical Model to Explain the Adoption of Continuous Cloud Service Certification by Certification Authorities, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1465-1476

ning up new resources when demand increases. Certifications in general are well-recognized means for organizations to assess goods and services [1, 2], and their importance and number steadily increase in recent years [3].

Yet, current research has primarily focused on identifying and assessing the effectiveness of certifications at a given point in time, and thus they are essentially regarded as static snapshots of attributes of providers and their services [4–6]. However, cloud service environments are highly dynamic, resulting from challenging cloud computing characteristics (e.g., on-demand provisioning and entangled supply chains), fast technology life cycles and ongoing architectural changes [6–8]. Likewise, cloud services are faced with dynamically emerging environmental challenges and with changes in legal landscape which might threaten certification effectiveness and reliability in the medium to long term.

We believe that current certifications fail to cope with an ever-changing cloud environment because certification assessments are based only on static, manual expert assessments and periodic spot checks, and may not be actually valid for longer periods of time since certification. Therefore, we argue that continuous service certification (CSC) is required to assure reliable and trustworthy certifications and cloud services.

CSC is beneficial for cloud certification authorities, service providers and customers altogether [7, 8]: certification authorities can actively detect and investigate critical certification deviations as they occur, thus increasing certification reliability over today's approaches; cloud providers can constantly improve their cloud services by evaluating ongoing feedback from certification authority about their performance; and finally CSC can counteract customers' worries due to lack of control of cloud infrastructure by increasing the transparency of providers' operation. With increasing reliance of organizations on cloud service providers, the necessity for continuous reliable, trustworthy and meaningful certification gains importance. Yet, CSC remains currently underexplored, not well test marketed and evaluated only in trials, resulting in a low adoption rate by certification authorities [8].

To understand and enhance CSC's rate of adoption, and therefore ultimately pave the way for continuously reliable and secure cloud services, we examine the adoption process of CSC from the perspective of certification authorities by conducting a thorough literature search, and building on the Diffusion of Innovations theory [9] and the Technology-Organization-Environment framework [10, 11]. Our work helps to answer the research question: *What influences certification authorities to adopt CSC?*

Investigating how the characteristics of an innovation as well as organizational and environmental factors will affect CSC's rate of adoption can be of great value to understand and enhance actual adoption processes [9]. With this study, we advance the understanding of the CSC adoption process by providing a synthesis and discussion of relevant factors that influence adoption rate of certification authorities. In addition, we provide a theoretical model to be tested in future research for validation.

The paper proceeds as follows. We provide a background on cloud service certifications and highlight the need for CSC, followed by a brief presentation of our research approach. Thereafter, we discuss how the characteristics of CSC, organizational, environmental, risk and cost factors influence the adoption process. We then discuss our findings and conclude with directions for future research.

2 Theoretical Background

2.1 Cloud Service Certifications

Cloud computing offers ubiquitous, on-demand access to a shared pool of configurable IT resources (e.g., servers, storage and applications) that can be rapidly provisioned and released with minimal management effort or service provider interaction [12]. On the one hand, cloud services offer an attractive alternative to traditional IT usage for organizations, and on the other hand they challenge contemporary security and privacy risk assessment approaches. Therefore, cloud services face a broad range of risks including lack of accessibility, reliability and virtualization vulnerabilities, privacy and control issues as well as issues related to data integrity and segregation [13].

One widespread strategy to reduce customers' uncertainties is to adopt certifications, which is particularly important for small and medium-sized cloud providers [1]. A certification is defined as a third party attestation of products, processes, systems or persons that verifies the conformity to specified requirements [14]. During a certification process, independent and accredited auditors perform comprehensive, manual checks to test adherence according to a defined set of certification criteria. If a provider adheres to the specified requirements, the certification authority awards a formal written certificate. A variety of certifications has already been developed and market tested to signal that providers have adopted their standards and comply with their certification audits; these exist particularly in cloud markets (e.g., EuroCloud *'StarAudit'* and Cloud Security Alliance *'Security, Trust & Assurance Registry'*). Cloud service certifications typically consist of security, privacy and reliability requirements, and build on IT standards (e.g., ISO 27001, ISO 27017 and ITIL), and aim to ensure availability, integrity and confidentiality of cloud services for a validity period of one to three years [15].

2.2 The Need for Continuous Certification

Existing certifications represent only a retrospective look at the fulfillment of technical and organizational measures. Requirements of certifications may no longer be met throughout the validity period of the certification because cloud services are confronted with continuously emerging environmental dynamics. Especially, we refer to environmental dynamics that are difficult to predict, lead to an instability and create uncertainty for customers or providers [16]. The premise behind these assumptions is that external environments impact organizational performance, and organizations must take into account environmental characteristics and emerging dynamics when formulating strategies and structures as well as during daily operations. As such, inherent cloud computing characteristics, ongoing architectural changes, the emergence of environmental threats or changes in legal and regulatory landscape can be regarded as dynamics that might have an impact on actions taken by a provider. Certification reliability has to be re-evaluated over time if the assumptions under which a certification was awarded have changed. Consequently, we believe that CSC is required to assure continuously reliable and trustworthy certification and cloud services. CSC is a methodology that enables certification authorities to react and to adjust their certification reports simultaneously with the occurrence of events concerning the cloud service [6].

3 Research Approach

3.1 Literature Analysis

In this study, we focus on identifying factors that influence the adoption of CSC by certification authorities, and therefore conducted a thorough literature review. To find pertinent literature that deals with innovation adoption processes, we performed a search in the online database of EBSCOHost (Academic Search Complete and Business Source Complete). This search was executed on 15th March 2016 and was based on the following search string: *("Diffusion of Innovation*") AND (("Information System*") OR (IS))*, inspired by the Diffusion of Innovations theory [9]. The search was limited to title, abstract and keywords. Moreover, the results were reduced by applying the filters for only "peer-reviewed publications". This initial search revealed 81 potentially relevant articles, published from 1982 to 2015, which deal in different ways with the adoption or the diffusion of innovations. Some of these publications deal with innovations in general whereas others specifically refer to concrete innovations. By examining these articles, we determined 55 of them suggesting factors influencing the adoption of an innovation. Identified papers were read and factors impacting the adoption of an innovation were marked for further analysis, despite individual findings relating to the factors (i.e., regarding their empirical support) [17]; leading to 437 factors. As a lot of different factors were used in different articles, sometimes under different name, but we named them only once and noted their frequency of being mentioned, as this can be seen as an indicator of their importance. By this we reduced our list of factors to 258.

To further reduce this number of factors we carefully analyzed the used terms and their meaning. First, we identified synonyms, aggregated them into one factor and summed up the frequency of being mentioned for each of the synonymous terms. Second, we subsumed terms with similar meanings as for example *"competitors"*, *"competition"*, *"competitive advantage"*, and *"other industry players"* to *"competitive pressure"* and considered their total frequency of mention. Third, we excluded terms which, for example, refer to the adoption process itself rather than to factors influencing the adoption decision like *"earliness of adoption"*, and those terms which are referring to a concrete innovation, for example, *"website features"* and therefore cannot be transferred to CSC context. Finally, we carefully analyzed whether remaining factors are empirically supported and read research findings to ensure relevancy of factors. Based upon the remaining factors and on the frequency of being mentioned, we formed five groups of factors which have a major influence on the adoption of an innovation: *innovation's characteristics* (mentioned 124 times) including relative advantage, complexity, compatibility, observability and trialability; *organizational factors* (66) including organization, management and technology attributes; *individual factors* (49) including attitudes and skills; *environmental factors* (34) including the legal and regulatory landscape, market and competitive pressure; and finally *risks and costs* (16).

3.2 Theories of Factors Influencing the Adoption of Continuous Certification

The five groups of factors resulting from our literature analysis are in line with and can be assigned to two different theoretical models explaining the adoption of innovations: the Diffusion of Innovations theory (DOI) and the Technology-Organization-Environment (TOE) framework. The DOI theory was proposed by Everett M. Rogers [9] and focuses on why innovations – although having obvious advantages – are often very hesitantly adopted. A central concept of the DOI theory is the diffusion process, in which an innovation is communicated through certain channels, over time, among the members of a social system. Information about the innovation will be communicated during the diffusion process, which reduces uncertainty of potential adopters about the innovation itself, and finally leads to an adoption or rejection decision. An innovation is defined as an idea, practice or object that is perceived as new by an individual. The adoption rate is defined as the relative speed with which members of a social system adopt an innovation. While most research has concentrated on the adoption of innovations in regard to differences in their innovativeness, DOI theory examines the innovation itself, and how its characteristics affect its rate of adoption. DOI theory describes five main innovation characteristics: *relative advantage*, *compatibility*, *complexity*, *trialability* and *observability*. Literature shows that the DOI theory has a solid theoretical foundation and consistent empirical support (e.g., [18–20]). DOI theory focuses on the impact of innovation's characteristics, but acknowledges that the specific context, for example, the organization and her environment can influence the adoption rate as well [9, 21]. We integrate the TOE framework that serves as an important, additional theoretical perspective for studying such contextual factors [10, 11].

The TOE framework was developed by De Pietro, Wiarda and Fleischer [10], and is embedded into the research by Tornatzky and Fleischer [11] who describe the entire process of technological innovation, from the invention or development by engineers until the adoption and implementation by users within an organization. The TOE framework focuses on factors that influence the adoption and implementation of innovations in the context of an organization. It identifies three main contexts that influence the adoption of innovation: the technological, organizational and environmental context [10, 11]. The TOE framework has been used by researchers to examine the adoption of technological innovations, and has received ample empirical support (e.g., [20, 22, 23]).

To construct our theoretical model, we combined the DOI theory and the TOE framework by using the *innovation's characteristics* as representative factors for the *technological context.* In addition, we considered *organizational* factors, including both managerial and IT capabilities. We complemented them by *environmental factors* as well as the factor group '*risks and costs'* because they take a decisive influence with regard to the adoption of CSC (see Figure 1). We excluded the group of *'individual'* factors - although resulting from the literature research - because this study takes an organization level perspective. Finally, we excluded *trialability* as one of the innovation's characteristics because CSC cannot be tested easily beforehand as it affords high efforts and expenditures.

Figure 1. Factors influencing the adoption of CSC.

4 Theoretical Model of Factors and their Impact on the Adoption of Continuous Certification

In the following, we discuss identified factors in regard to the adoption of CSC by certification authorities, derive propositions about their impact and integrate them into a theoretical model.

Innovation's characteristics exert a great influence on the adoption of an innovation. Before an organization passes through the innovation-decision process, it seeks information in order to decrease uncertainty about the relative advantage of an innovation [9]. Such a *relative advantage,* which for example generates cost savings or offers the solution to an existing problem, can lead to the adoption of an innovation because it is perceived as better, more economic or expediently. Providing CSC services is beneficial for certification authorities because CSC increases their efficiency and reliability of issued certifications in particular.

In the context of traditional certification processes, adherence to certification requirements is observed by spot checks on a yearly basis only. Hence, certification deviations might be detected lately or hardly ever. In contrast, CSC allows the certification authority to actively detect critical defects as they occur. Hence, CSC can be considered as proactive and enables corrective actions as soon as a problem is detected. So CSC can improve reliability and trustworthiness of issued certifications. In addition, certification reports are more relevant to customer's decision makers. The change from yearly spot checks to CSC is often accompanied by the use of automated certification processes which enable certification authorities to test larger data samples and examine data in a faster and therefore more efficient way, compared to their manual predecessors. Finally, the certification authority might gain further benefits by offering innovative certification services for cloud customers and charging extra fees (e.g., enabling customers to validate requirement adherence on demand). While in traditional certification contexts a business relationship only exists between the cloud provider and the certification authority, CSC enables certification authorities to build up a direct relationship with cloud service customers, hence, creating new business models. Consequently, CSC provides significant relative advantages for certification authorities because it increases the efficiency and quality of certifications, enables new business models, and leads to continuously secure and reliable cloud services.

Proposition 1 (P1): Relative advantages foster the adoption of CSC by certification authorities.

The more *compatible* an innovation is perceived with sociocultural values and beliefs, the needs of potential adopters or with previous experiences the less uncertainty concerning the innovation is present; leading to a higher rate of adoption [9]. Certifications are well-recognized means for customers to assess goods and services [1]. Importance and number of independent third party product and service assessments steadily increase in recent years [3]. Yet, providers are threatened by a highly dynamic and ever-changing environment, and thus quickly respond to emerging environmental dynamics. With increasing reliance of customers on cloud services their demand for continuous, highly reliable and secure services gains importance. Consequently, it is necessary for the certification authority to continuously verify the conformity with certification requirements.

Previously introduced ideas and practices are a familiar standard against which the innovation can be interpreted [9]. Current certification practices are mostly based upon manual auditing operations, for example, performing interviews and manual security tests. The transition to CSC requires an automation of certification processes. The use of computer-based audit tools and technologies (CAATTs), which already aims at automating processes and facilitating the certification authority's work, could therefore promote this transition. Nonetheless, surveys reveal that CAATTs are not yet frequently and systematically used [24], although they are seen as useful and beneficial. We assume that CSC is compatible with the needs of relevant stakeholders and previously introduced ideas leading to a positive effect on the adoption.

P2: A high compatibility fosters the adoption of CSC by certification authorities.

The *complexity* of an innovation is measured by the degree to which the innovation is perceived as relatively difficult to understand and use [9]. The higher the complexity is, the greater is the uncertainty of potential adopters. Adopting CSC exhibits a high degree of complexity. Certification authorities must establish CSC and management systems to support the certification planning, management, operation and scheduling activities, develop new certification processes and train their employees. In order to reduce the complexity of the CSC, authorities can build on existing monitoring systems and processes of the provider to gather certification-relevant data [25]. For example, certification authorities might access an interface that enables the secure and reliable transmission of relevant data. Further on, the authority has not only to manage his own CSC operations, but also has to consider and align with providers' ongoing activities, which also increases the complexity of CSC. Consequently, the certification scope has to be adjusted individually for each cloud service, for example, in regard to available cloud systems, provider's organizational size, the number of employees as well as the level of technical knowledge and skills.

P3: A high complexity hampers the adoption of CSC by certification authorities.

The *observability* is the degree to which the innovation provides tangible results [9]. The higher the perceived observability of an innovation is, the more positively it affects the adoption rate. Performing CSC aims to increase transparency about cloud service operation and certification adherence. Results of CSC will be visible for the public, for example, by ongoing certification reports. In order to further increase the observability, CSC offers the means for a new generation of web assurance seals: dynamic, up-to-date, and accurate seals informing customers about the actual certification requirement adherence status. Creating a high transparency for cloud service customers promotes the observability of CSC and has a positive effect on its adoption rate. However, a high observability also places high burdens on the protection and anonymization of provided data to ensure data confidentiality, integrity and authenticity.

P4: A high observability fosters the adoption of CSC by certification authorities.

Organizational factors comprise features and characteristics of the organization, essential aspects of management as well as the extent and the level of use of technology; factors that influence the adoption of an innovation [10, 11]. An organization is characterized by its *age* and *size* among others. Since size represents several important aspects of an organization, such as slack resources, organizational structure and decision-making flexibility, it is a critical factor to influence innovation adoption [9]. In the IS literature different opinions exist regarding the role that size plays [20]. On the one hand, large and established authorities may be less flexible than smaller and younger organizations, might show less innovation readiness and rather insist on previously applied methods [20, 23]. But on the other hand, these authorities have access to profound experience and knowledge about certification processes and emerging innovations, and can build on more financial means and multifarious human and material resources in order to meet challenges posed by the adoption of innovations [9, 20]. We assume that certification authority's size and age will foster the adoption of the CSC because they generally possess slack resources and expertise to meet adoption challenges, including high initial investments and the redesign of certification business processes.

P5: The certification authority's size and age will foster the adoption of the CSC.

With respect to organization's *management*, its settings, policies and priorities in particular affect the adoption of innovations [10, 11]. Thus, for example, CSC adoption should be consistent with organizational objectives and strategy [21] and supported by the top management [26]. The top management should provide the vision, support and commitment around the innovation as well as commit resources and create the environment required for the adoption [27]. Thus, top management exerts a positive influence on the adoption of CSC.

P6: Management support fosters the adoption of CSC by certification authorities.

Further on, the certification authority's *technology competence* has an influence on the innovation adoption [10, 11, 20]. Technology competence refers to the technological characteristics available in the organization such as the IT infrastructure and IT professionals [23]. The IT infrastructure covers the installed technologies, systems and applications within the certification authority allowing an integration of CSC services and corresponding IT systems. IT professionals are the human resources with technical knowledge required to efficiently perform CSC. For example, if the existing IT infrastructure is highly developed and versatile, and supports the integration of new CSC components, adoption uncertainty is reduced and adoption rate increases.

P7: Technology competence fosters the adoption of CSC by certification authorities.

Environmental factors comprise environmental values and norms, customer and competitive pressures [10, 11]. The *environmental values and norms* can affect the adoption of CSC for example by changing or setting up new guidelines. If for example the validity period of cloud certificates is generally shortened, this prepares the transi-

tion to the CSC and could ultimately effect that certification authorities are only accredited when awarding their certificates based on CSC. Also the government can contribute to the adoption of CSC when well-reputed government institutions highlight the use of CSC as an effective way to increase the security and reliability of cloud services.
P8: The values and norms foster the adoption of CSC by certification authorities.

Cloud customer pressure can exert great influence on the adoption of CSC [9–11]. Certification authorities might start adopting CSC, for example, if an ever-increasing amount of (potential) cloud customers demands reliable certifications in modern, turbulent environments. In the future, customers might decide whether to use a cloud service or not, based on providers' willingness to be continuously certified. Consequently, customer pressure is assumed to be of great influence on the adoption of CSC.
P9: Customer pressure fosters adoption of CSC by certification authorities.

Certification authorities compete for certification requests. *Competitive pressure* also acts as a facilitator influencing the adoption of CSC. Either the incentive of first mover competitive advantages or the urgency to keep up with competitors will provide the focus and purpose to successfully overcome obstacles and resistance to innovation adoption within an organization [21, 28]. Likewise, innovation imposition strategies by providers and partners might foster adoption rate of CSC, for example, if cloud providers tend to engage only with certification authorities that apply CSC in order to fulfill the demands of their cloud customers. Subsequently, competitive pressure might force certification authorities to open up for CSC and to create necessary conditions for adoption.
P10: Competitive pressure fosters the adoption of CSC by certification authorities.

Risk and cost factors are referring to possible disadvantages or dangers, and to expenses, which may affect the adoption of an innovation. In general, risks and costs represent multi-dimensional constructs that need to be viewed from different angles and analyzed in detail. For example, various security and privacy risks might emerge that impact certification authorities' adoption intention differently. CSC implies the transmission and storage of data about the cloud service at the site of the certification authority. Subsequently, certification authorities are becoming a valuable target of attackers from the outside. Hence, this involves high risks of data theft, leads to significantly higher demands on data security and data protection, and may hamper the adoption.
P11: Risks hamper the adoption of CSC by certification authorities.

CSC of cloud services usually goes with automation of processes which on the one hand affords high expenditures for purchasing the technical equipment and a high amount of running costs for the operation and maintenance. On the other hand, an automation of processes might lead to (mid-term) cost savings.
P12: Costs influence the adoption of CSC by certification authorities.

Figure 2 depicts our theoretical model and summarizes identified factors and their impact on the adoption intention of CSC by certification authorities. Adoption intention refers to the probability that an organization will adopt CSC processes, set up required IT infrastructures, and provide CSC services for cloud providers.

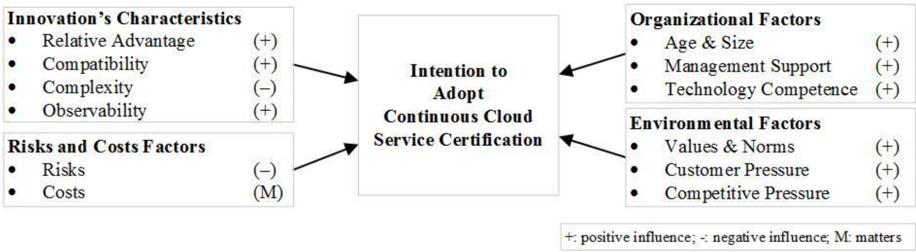


Figure 2. Theoretical model of CSC adoption by certification authorities.

5 Discussion and Conclusion

Based on a literature analysis, we developed a theoretical model by integrating the DOI theory and the TOE framework complemented by risk and cost factors to examine which factors influence the adoption of continuous cloud service certification. Thereby, we are able to analyze the adoption of an innovation from two different perspectives: the innovation itself with its characteristics and the surrounding organizational and environmental contexts. This study shows that many factors have an important impact on the adoption of CSC. We believe that the multifarious relative advantages of CSC and a high degree of observability will strongly motivate certification authorities to adopt CSC. On the other hand, a limited compatibility and a high complexity might hamper adoption. In regard to organizational factors, top management support and a high technical competence will positively influence the adoption of CSC. As environmental factors, customer and competitive pressures are of great importance when adopting CSC of cloud services. At last, risks and costs are relevant inhibitors for the adoption of CSC.

The identified and discussed factors have been considered separately, but some are closely related to each other, which might result in moderating effects on the adoption intention. First, relative advantages of CSC – due to a high observability of CSC results – are visible for both cloud customers and competitors, and thus they can lead to an increase of customer and competitive pressure as environmental factors. Second, a high technological competence, for example, due to the existence of a well-equipped IT department with well-trained specialists, reduces the complexity of CSC as well as increases innovation's compatibility. Finally, a close interrelationship between environmental pressures as well as perceived relative advantages, and top management support is apparent because they influence the strategy of an organization and actions that are preferred by the management.

With this study, we provide a two-fold contribution for research and practice. First, we advance the understanding of the CSC adoption process by providing a synthesis and discussion of relevant factors that influence adoption rate from a DOI and TOE perspective. Investigating how the attributes of an innovation affect its rate of adoption can be of great value to change agents seeking to predict the reactions to an innovation, and perhaps to modify certain of these reactions by the way they name and position an innovation [9]. Finally, we provide a theoretical model to be tested in future research to validate our assumptions, and to enhance the adoption process.

Nevertheless, this study has some limitations. Our discussion of the factors is based on literature analysis and theoretical reasoning research only since at the current diffusion state only a minority of certification authorities have started to deal with CSC adoption. However, we are currently working on a quantitative study to analyze to what extent the discussed factors influence CSC adoption. Within this study we focused on the adoption of CSC of cloud services by certification authorities, hence our theoretical model might be limited in regard to the context of cloud services as well as for the certification authorities as stakeholder. Finally, we neglected factors of individual adopters (i.e., managers) which might be of great importance in the actual adoption decision process.

"Last, […] an innovation's rate of adoption is affected by the extent of change agents' promotion efforts" [9]. On this account, we want to encourage researchers and practitioners with this study to participate in adopting and diffusing CSC.

6 Acknowledgements

This research is funded by the German Federal Ministry for Education and Research (grant no. 16KIS0079).

References

1. Sunyaev, A., Schneider, S.: Cloud services certification. CACM 56, 33–36 (2013)
2. Schneider, S., Sunyaev, A.: Determinant factors of cloud-sourcing decisions: reflecting on the IT outsourcing literature in the era of cloud computing. Journal of Information Technology 31, 1–31 (2016)
3. International Organization for Standardization: The ISO Survey of Management System Standard Certifications – 2014. Executive summary
4. Connelly, B.L., Certo, S.T., Ireland, R.D., Reutzel, C.R.: Signaling Theory. Journal of Management 37, 39–67 (2011)
5. Etzion, D., Pe'er, A.: Mixed signals. A dynamic analysis of warranty provision in the automotive industry, 1960-2008. Strategic Manage J 35, 1605–1625 (2014)
6. Lins, S., Schneider, S., Sunyaev, A.: Trust is Good, Control is Better. Creating Secure Clouds by Continuous Auditing. IEEE Transactions on Cloud Computing (2016)
7. Lins, S., Grochol, P., Schneider, S., Sunyaev, A.: Dynamic Certification of Cloud Services. Trust, but Verify! IEEE Security & Privacy 14, 66–71 (2016)
8. Lins, S., Teigeler, H., Sunyaev, A.: Towards a bright future: Enhancing diffusion of continuous cloud service auditing by third parties. In: Proceedings of the 24th European Conference on Information (2016)
9. Rogers, E.M.: Diffusion of innovations. Free Press, New York (1962)
10. DePietro, R., Wiarda, E., Fleischer, M.: The context for change. In: Tornatzky, L.G., Fleischer, M., Chakrabarti, A.K. (eds.) The processes of technological innovation. Lexington Books (1990)
11. Tornatzky, L.G., Fleischer, M., Chakrabarti, A.K. (eds.): The processes of technological innovation. Lexington Books (1990)

12. Mell, P.M., Grance, T.: The NIST definition of cloud computing. National Institute of Standards and Technology, Gaithersburg, MD (2011)
13. Subashini, S., Kavitha, V.: A survey on security issues in service delivery models of cloud computing. J Netw Comput Appl 34, 1–11 (2011)
14. International Organization for Standardization: Conformity assessment - Vocabulary and general principles 03.120.20; 01.040.03
15. Schneider, S., Lansing, J., Fangjian Gao, Sunyaev, A.: A Taxonomic Perspective on Certification Schemes. In: Proceedings of the 47th Hawaii International Conference on System Sciences, pp. 4998–5007
16. Miles, R.E., Snow, C.C., Pfeffer, J.: Organization-Environment: Concepts and Issues. Industrial Relations: A Journal of Economy and Society 13, 244–264 (1974)
17. Lacity, M.C., Khan, S., Yan, A., Willcocks, L.P.: A review of the IT outsourcing empirical literature and future research directions. Journal of Information Technology 25, 395–433 (2010)
18. Premkumar, G., Ramamurthy, K., Nilakanta, S.: Implementation of Electronic Data Interchange 11, 157–186 (1994)
19. Beatty, R.C., Shim, J.P., Jones, M.C.: Factors influencing corporate web site adoption. Information & Management 38, 337–354 (2001)
20. Zhu, K., Dong, S., Xu, S.X., Kraemer, K.L.: Innovation diffusion in global contexts. Eur J Inf Syst 15, 601–616 (2006)
21. Bradford, M., Florin, J.: Examining the role of innovation diffusion factors on the implementation success of enterprise resource planning systems. International Journal of Accounting Information Systems 4, 205–225 (2003)
22. Grover, V.: An Empirically Derived Model for the Adoption of Customer-based Interorganizational Systems. Decision Sciences 24, 603–640 (1993)
23. Zhu, K., Kraemer, K.L.: Post-Adoption Variations in Usage and Value of E-Business by Organizations. Inform Syst Res 16, 61–84 (2005)
24. Mahzan, N., Lymer, A.: Examining the adoption of computer-assisted audit tools and techniques. Managerial Auditing Journal 29, 327–349 (2014)
25. Stephanow, P., Fallenbeck, N.: Towards continuous certification of Infrastructure-as-a-service using low-level metrics. In: International Conference on Advanced and Trusted Computing, pp. 1–8 (2015)
26. Liang, H., Saraf, N., Hu, Q., Xue, Y.: Assimilation of Enterprise Systems. MIS Quarterly 31, 59–87 (2007)
27. Lee, S., Kim, K.-j.: Factors affecting the implementation success of Internet-based information systems. Computers in Human Behavior 23, 1853–1880 (2007)
28. Zaltman, G., Duncan, R., Holbek, J.: Innovations and organizations. Wiley, NY (1973)

Das Business Model House of Quality: Bewertung plattformbasierter Geschäftsmodelle mit Quality Function Deployment

Felix Schönhofen[1] and Sixten Schockert[1]

[1] University of Stuttgart, Chair of Information Systems II, Stuttgart, Germany
{schoenhofen,schockert}@wius.bwi.uni-stuttgart.de

Abstract. Companies in the IT industry are exposed to permanent change. These changes relate in particular to emerging markets and business models. To ensure the company's long-term success, their business models have to be evaluated continuously. Due to the high amount of different stakeholders in multisided markets, this challenge becomes even greater for companies which business models are based on multisided platforms. The following paper shows the possibility to evaluate business models in multisided markets using Quality Function Deployment (QFD). In cooperation with companies in the sharing economy QFD workshops were held to reach this goal and to evaluate the QFD approach to examine business models in multisided markets.

Keywords: Business Models, Platforms, Quality Function Deployment, Requirements, Multisided Markets

1 Motivation

Der fortschreitende Trend der Digitalisierung stellt einerseits große Herausforderungen an Unternehmen, verspricht andererseits aber auch Potenzial für die Entwicklung neuartiger Geschäftsmodelle [1]. Eine Reihe bekannter Neugründungen wie Uber, Car2go und Airbnb haben dies in den letzten Jahren erfolgreich ausnutzen können [2].

Diesen Unternehmen ist gemein, dass der Kern ihres Geschäftsmodells der Ausgangspunkt eines sogenannten Wertschöpfungsnetzwerks ist [3]. Ihr Produkt ist entsprechend eine Plattform, die einer größeren Bandbreite an Kundenanforderungen begegnen kann, als es dem Produkt eines einzelnen Anbieters möglich wäre [4].

Den genannten positiven Beispielen steht jedoch eine deutlich größere Anzahl gescheiterter Startups gegenüber [5]. Häufig sind eine fehlende Marktnachfrage und ein unvollständiges oder gar fehlendes Geschäftsmodell Grund für den Misserfolg [6]. Offensichtlich ist neben der richtigen Produktidee die Erarbeitung eines tragfähigen Geschäftsmodells eine wichtige Voraussetzung für einen erfolgreichen Markteintritt.

Die Auswirkungen des dynamischen Umfelds in dem die Unternehmen agieren, der hohe Grad an Unsicherheit aufgrund der Neuartigkeit der Produkte und Dienstleistungen sowie der Einfluss der Informations- und Kommunikationstechnik auf die Geschäftsmodelle erfordert von den Unternehmen darüber hinaus die Fähigkeit, sich

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Schönhofen, F.; Schockert, S. (2017): Das Business Model House of Quality: Bewertung plattformbasierter Geschäftsmodelle mit Quality Function Deployment, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1477-1488

schnell an ändernde Umgebungen anpassen zu können [7-8]. Hierfür ist allerdings eine Bewertung, ob und in welcher Ausprägung Änderungsbedarf besteht unerlässlich.

Im vorliegenden Beitrag soll ein Verfahren vorgeschlagen werden, wie plattformbasierte Geschäftsmodelle untersucht und hinsichtlich möglicher Schwachstellen bewertet werden können. Die Grundidee dabei ist, dass die konkreten Ausgestaltungen eines plattformbasierten Geschäftsmodells nicht zufällig gewählt werden, sondern begründet sind, d. h. die Ausprägungen der Merkmale eines plattformbasierten Geschäftsmodells beruhen auf möglichen Anforderungen an eben dieses Geschäftsmodell. Diese Zweck-Mittel-Beziehungen werden in diesem Beitrag mittels der ursprünglich für Produkt(weiter)entwicklungen verwendeten Methode Quality Function Deployment (QFD) dargestellt und ermittelt [17], [24]. QFD eignet sich dafür in besonderen Maße, da es zum einen mit seinen methodischen Merkmalen überall dort anwendbar ist, wo gemeinschaftlich in einem Team Lösungen zu bestimmten Problemen bzw. Antworten auf Anforderungen gesucht werden. Zum anderen trennt QFD analytisch konsequent zwischen Anforderungen und möglichen Lösungen und schafft so das Potential für die Gestaltung von alternativen, besseren Lösungen [17].

Entsprechend sollen die folgenden Fragen beantwortet werden:

Tabelle 1. Untersuchungsziele des Beitrags

Forschungsfragen
1. Welche Merkmale weisen plattformbasierte Geschäftsmodelle auf?
2. Was sind Anforderungen an plattformbasierte Geschäftsmodelle?
3. Wie kann ein Vorgehensmodell zur Bewertung plattformbasierter Geschäftsmodelle auf Basis von QFD gestaltet sein?

Zur Beantwortung dieser Fragen folgt dieser Beitrag grundsätzlich dem gestaltungsorientierten Forschungsansatz gemäß des Design Science Ansatzes [25]. Der Aufbau des Beitrags folgt entsprechend dem Publication Schema für Design Science Research von Gregor und Hevner [26].

2 Stand der Forschung

Seit den 1990er Jahren findet der Begriff Geschäftsmodell in der Öffentlichkeit zunehmend Beachtung und wird sowohl in Wissenschaft als auch Praxis häufig verwendet [9-10]. Trotz dieser Verbreitung existiert bis heute keine einheitliche bzw. allgemein akzeptierte Definition dieses Begriffes [11].

Eine häufig identifizierte Gemeinsamkeit der unterschiedlichen Definitionen ist die Position des Geschäftsmodells als Intermediär zwischen der Ebene der Geschäftsprozesse und der Unternehmensstrategie [1], [12-13]. Eine entsprechende Definition stammt von Al-Debei et al.: „A business model is a conceptual tool that contains a set of elements and their relationships and allows expressing a company's logic of earning money. It is a description of the value a company offers to one or several segments of customers and the architecture of the firm and its network of partners for creating, marketing and delivering this value and relationship capital, in order to generate profitable and sustainable revenue streams.“ [13].

In dem vorliegenden Beitrag sollen insbesondere solche Geschäftsmodelle betrachtet werden, die Ausgangspunkt eines Wertschöpfungsnetzwerkes sind. Solche Ausgangspunkte können sogenannte Plattformen gemäß der Definition von Branchenplattformen nach Gawer sein: Diese Plattformen umfassen alle Akteure, die innerhalb einer Branche zusammenarbeiten: Der Anbieter der Plattform muss zum Ziel haben nicht nur von den eigenen, sondern auch von komplementären Innovationen zu profitieren. Die Komplementoren selbst müssen versuchen von indirekten und direkten Netzeffekten zu profitieren [14]. Direkte Netzeffekte bezeichnen die positive Korrelation aus dem Nutzen den der einzelne Akteur aus dem Netzwerk zieht und der Gesamtzahl der Akteure. Die Indirekten werden durch die Größe der jeweilig anderen Marktseite bestimmt: Eine größere Anzahl an auf der Plattform versammelten Endkunden verspricht für die Komplementoren einen größeren Markt und vergrößert entsprechend den Anreiz Komplemente zu entwickeln. Umgekehrt bedeutet eine größere Zahl an Komplementen eine höhere Attraktivität der Plattform für die Endkunden [3], [8].

Aus diesen Zusammenhängen konstituiert sich die Grundherausforderung bei der Etablierung eines plattformbasierten Geschäftsmodells. Es muss eine Lösung für das sogenannte Henne-Ei-Problem gefunden werden: Ohne eine bedeutende Zahl an Endkunden, werden die Komplementoren der Plattform fernbleiben und umgekehrt [15]. Die Bedeutung dieses Kernproblems lässt sich an bekannten Beispielen wie BlackBerry beobachten.

Weitere Herausforderungen treten ein, wenn das Henne-Ei-Problem gelöst ist: Die eigentlich kooperierenden Akteure innerhalb des Netzwerkes können gleichzeitig in Konkurrenz zueinanderstehen. In der Literatur hat sich für diese Problematik der Begriff Koopkurrenz durchgesetzt [16]. Im Beispiel eines Vermittlers von Fahrgemeinschaften wird das Problem offensichtlich: Aufgrund der indirekten Netzeffekte profitiert der Anbieter einer Fahrt von einer hohen Zahl an anderen Fahrtanbietern und einem entsprechend hohen Angebot an Fahrten auf der Plattform. Allerdings erhöht eine hohe Zahl an Fahrtanbietern auch die Wahrscheinlichkeit, dass Fahrten angeboten werden, die in direkter Konkurrenz zu seinem eigenen Fahrtangebot stehen. Für den einzelnen Anbieter entsteht so die Konstellation, dass er gleichzeitig von den anderen Anbietern profitiert und mit ihnen konkurriert.

Beide Herausforderungen rechtfertigen so eine gezielte Betrachtung von diesen Geschäftsmodellen in mehrseitigen Märkten, die in diesem Beitrag mit plattformbasierten Geschäftsmodellen bezeichnet wird.

3 Untersuchungsdesign

Im folgenden Kapitel wird dargestellt, wie Anforderungen an plattformbasierte Geschäftsmodelle ausgehend von bestehenden Geschäftsmodellmerkmalen erarbeitet werden können. Hiermit sollen die Forschungsfragen 1 und 2 adressiert werden.

3.1 Quality Function Deployment

Im Rahmen der methodischen Produktentwicklung hat sich die Methode *Quality Function Deployment* (QFD) seit den 1960er Jahren bewährt [17], [24]. QFD wird definiert als „[....] method for structured product planning and development, that enables a development team to specify clearly the customer's wants and needs, and then to evaluate each proposed product or service capability systematically in terms of its impact on meeting those needs". Die zugrundeliegenden methodischen Merkmale von QFD lassen allerdings eine Anwendung auch auf andere Disziplinen zu [18]. Diese Vielseitigkeit ermöglicht es, QFD auch als Methode zur Bewertung plattformbasierter Geschäftsmodelle zu nutzen.

Als Vorteile wird von Herzwurm et al. unter anderem die Schaffung einer hohen Kundenzufriedenheit genannt, da sich die Entwicklungsvorgaben an den tatsächlichen Kundenbedürfnissen orientieren. Mit *Kunden* sind hierbei jedoch nicht (nur) die Kunden im Sinne möglichen Käufer eines Produktes gemeint, sondern alle Akteure (Stakeholder), die ein Interesse an dem zu entwickelnden Objekt haben [17]. In diesem Gedanken findet sich auch die Möglichkeit zur Anwendung von QFD für Geschäftsmodelle wieder: Auch wenn es sich bei einem Geschäftsmodell nicht um ein konkretes Produkt handelt, gibt es dennoch Stakeholder, die ein Interesse an einer möglichst erfolgsversprechenden Ausgestaltung haben.

Eines der charakteristischen Grundmerkmale von QFD ist die „explizite Trennung von Anforderungen und Lösungen", weiterhin sollen priorisierte Ziele formuliert und entlang des gesamten Entwicklungsprozesses verfolgt werden. Die genannten Anforderungen und Lösungen werden im sogenannten *House of Quality* (HoQ, siehe Abbildung 1) dargestellt. Neben den Anforderungen (*WAS*) und den Maßnahmen zur Anforderungserfüllung (*WIE*) enthält das HoQ die Korrelationen zwischen *WAS* und *WIE* [17]. Das grau hinterlegte „Dach" des HoQ wird in diesem Beitrag nicht betrachtet, bietet aber Anknüpfungspunkte für weitere Untersuchungen, siehe dazu Kapitel 5.

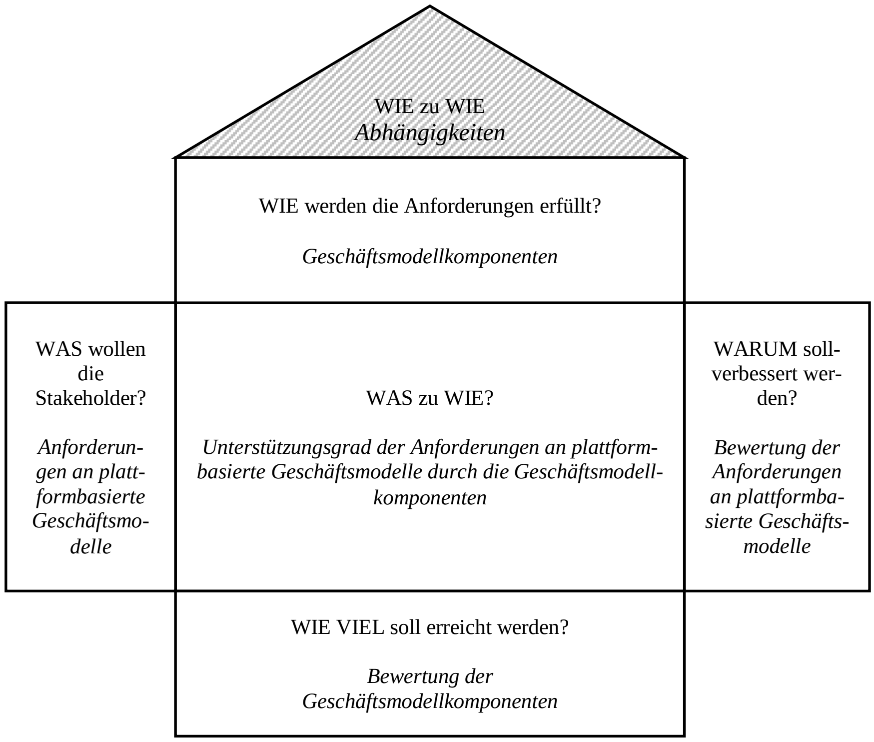


Abbildung 1. House of Quality plattformbasierter Geschäftsmodelle [17-18]

Grundidee ist, dass die Anforderungen mit den Lösungen auf eine bestimmte Art und Weise verknüpft sein müssen: Lösungen denen keine Anforderungen gegenüberstehen wären – eine korrekte und vollständige Exploration der Anforderungen vorausgesetzt – überflüssig oder würden zumindest keinen Beitrag zur Erhöhung der Kundenzufriedenheit beitragen. Umgekehrt und auf die Untersuchung von Geschäftsmodellen adaptiert bedeutet dies Folgendes: Weist ein Geschäftsmodell bestimmte Merkmale auf, muss es hierfür Gründe bzw. Anforderungen geben, die den das Geschäftsmodell erstellenden Akteur dazu veranlasst haben, dieses Merkmal zu berücksichtigen. Die Frage nach diesen Gründen könnte Hinweise auf die Anforderungen geben.

3.2 Vorgehen zur Identifikation der Anforderungen

Dieser Zusammenhang soll im Folgenden genutzt werden, um Anforderungen an plattformbasierte Geschäftsmodelle zu identifizieren. Die Erarbeitung dieser erfolgte in zwei Workshops in Zusammenarbeit mit Unternehmen aus dem Bereich der Sharing Economy, da hier die Vermittlungsleistung häufig im Vordergrund steht und entsprechende Plattformen häufig Kernbestandteil der Geschäftsmodelle von Unternehmen sind [19].

Mit Hilfe der sogenannten *Business Model Canvas* von Osterwalder und Pigneur [20] wurden während der Workshops Vertreter der Unternehmen nach der Ausgestaltung ihrer Geschäftsmodelle gefragt. Die Einteilung eines Geschäftsmodells in neun Kategorien wurde gewählt, da dieses Framework einen hohen Bekanntheitsgrad bei den untersuchten Unternehmen aufweist und entsprechend geringerer Erläuterungsbedarf bestanden hat [21]. Im Anschluss an diesen ersten Schritt wurden in einer moderierten Gruppendiskussion die Gründe für die jeweilige Ausgestaltung der Geschäftsmodellkomponenten erfragt und die Ergebnisse dokumentiert. Entsprechend der obigen Logik konnten dann in der Nachbereitung aus diesen Gründen Anforderungen an plattformbasierte Geschäftsmodelle extrahiert werden.

Mit diesem Schritt ist die erste Iteration des Vorgehens abgeschlossen. Die erarbeiteten Anforderungen dienen dann als Ausgangspunkt für die folgenden Workshopiterationen mit weiteren Unternehmen: Die Anforderungen werden gemäß Abbildung 1 als Anforderungen in das House of Quality eingetragen und können so den Geschäftsmodellmerkmalen der weiteren Workshoppartner gegenübergestellt werden. Der Unterstützungsgrad der Anforderungen durch die Geschäftsmodellkomponenten wird in der HoQ-Korrelationsmatrix folgendermaßen festgelegt [17-18]:

- 0: Geschäftsmodellmerkmal leistet keinen Beitrag zur Erfüllung der Anforderung
- 3: Geschäftsmodellmerkmal leistet schwachen Beitrag zur Erfüllung der Anforderung
- 9: Geschäftsmodellmerkmal leistet starken Beitrag zur Erfüllung der Anforderung

Ist dieser Vorgang abgeschlossen, kann die Korrelationsmatrix analysiert werden. Von Interesse ist hier insbesondere die Frage, ob sich sogenannte *schwache Zeilen* oder *Spalten* ergeben. Hiermit sind Zeilen/Spalten gemeint, die durchgängig niedrige Korrelationen aufweisen (also überwiegend den Wert 0 und nur wenige den Wert 3 enthalten). Eine solche schwache Zeile bedeutet, dass keines der Geschäftsmodellmerkmale (bzw. Ausprägungen von diesen) einen starken Beitrag zur Erfüllung der in dieser Zeile liegenden Anforderung leistet. Weiteren Aufschluss über die unterschiedliche Bewertung von Kundenanforderungen gibt das sogenannte *Kano-Modell zur Kundenzufriedenheitsmessung* [22]. Dementsprechend existieren Merkmale, die der Kunde nicht explizit verlangt, die bei Nichterfüllung allerdings Unzufriedenheit erzeugen (*Basismerkmale*). Merkmale, die der Kunde nicht verlangt, die allerdings bei Erfüllung die Zufriedenheit erhöhen und bei Nichterfüllung keine Unzufriedenheit auslösen, werden als *Begeisterungsmerkmale* bezeichnet. Merkmale, die bei Erfüllung Zufriedenheit und bei Nichterfüllung Unzufriedenheit auslösen, werden *Leistungsmerkmale* genannt. Diese Erkenntnis kann bei der Erarbeitung von Anforderungen über die Abfrage von Gründen für die Ausgestaltung des Geschäftsmodells von Bedeutung sein: Basis- und Begeisterungsmerkmale können Lösungen sein, zu denen (fehlerhafterweise) keine explizite Anforderung formuliert wurden [22].

Entsprechend diesem Modell können mögliche Gründe für eine solche schwache Zeile sein:

1. Die Anforderung ist von den Stakeholdern niedrig priorisiert und entsprechend werden keine Ressourcen für die Erfüllung dieser Anforderung verwendet.

2. Das verwendete Framework der Geschäftsmodelleinteilung weist eine Lücke auf.
3. Die Ausprägung der Geschäftsmodelle löst die Anforderung unzureichend.

Eine schwache Spalte bedeutet analog, dass das Geschäftsmodell Lösungsmerkmale aufweist, denen keine Anforderung gegenübersteht. Gründe hierfür können sein:

4. Es handelt sich um ein Begeisterungsmerkmal, nach dem die Stakeholder nicht explizit fragen.
5. Es handelt sich um ein Basismerkmal, nach dem die Stakeholder nicht explizit fragen.
6. Die vom betrachteten Lösungsmerkmal adressierte Anforderung wurde bisher noch nicht identifiziert.
7. Das Lösungsmerkmal ist überflüssig.

Hinweise auf den zutreffenden Grund kann die Priorisierung der Anforderungen und die Bewertung der Lösungsmerkmale bieten. Hierzu werden die bis zu diesem Zeitpunkt erarbeiteten Anforderungen und Geschäftsmodellmerkmale stakeholderspezifisch priorisiert. Dies kann über eine weitere Befragung der Workshopteilnehmer realisiert werden. Aufgrund der überschaubaren Anzahl an Anforderungen kann hier der paarweise Vergleich genutzt werden. Die Ergebnisse der Priorisierung werden im Anschluss in das HoQ eingetragen und können dann folgendermaßen zur Untersuchung genutzt werden [17]:

- Ist die Anforderung durch mindestens eine Stakeholdergruppe nicht niedrig priorisiert, kann 1. nicht zutreffen und 2. und 3. sind mögliche Gründe.
- Ist entsprechend ein Lösungsmerkmal nicht niedrig bewertet, aber trotzdem Ausgangspunkt einer schwachen Spalte, kann das ein Hinweis auf 6. sein, möglicherweise aufgrund von 4. oder 5.
- Entsprechend kann ein niedrig bewertetes Lösungsmerkmal in schwacher Spalte auf das Vorliegen von 7. hindeuten.

Sollten sich Hinweise auf 2., 3. oder 6. andeuten, wird damit die Notwendigkeit eines weiteren Workshops evident. Dieser verläuft analog mit dem ersten, wird allerdings am Ende um einen Vergleich der im letzten Workshop (oder im Falle mehrerer, aller bis dahin durchgeführten) erzielten Artefakte erweitert. Dieser Vergleich kann zu weiterem Informationsgewinn führen und bereits bei der Anforderungserhebung für die Experten zu einem Wissensgewinn führen.

Nach einer bestimmten Anzahl an Workshopiterationen wird die Untersuchung beendet. Der Zeitpunkt dieses Abbruchs bemisst sich am Grenzinformationsgewinn, der ein gewisses Limit unterschreitet (also keinem Vorliegen mehr der o.g. Punkte 2., 3. oder 6.).

3.3 Merkmale plattformbasierter Geschäftsmodelle

Die bereits erwähnte Business Model Canvas diente während der Workshops als Ausgangspunkt für die Erarbeitung der gesuchten Anforderungen. Im Vergleich zu den in diesem Framework vorgeschlagenen neun Komponenten umfasst der während der

Workshops erarbeitete Merkmalskatalog für plattformbasierte Geschäftsmodelle elf Komponenten. Die zusätzlichen Komponenten beziehen sich auf die Integration der komplementären Anbieter. Auf die Komponente *Kostenstruktur* wurde aufgrund der zu erwartenden tautologischen Zusammenhänge mit den Anforderungen verzichtet.

Tabelle 2: Merkmale plattformbasierter Geschäftsmodelle [20]

Nr.	*Merkmale plattformbasierter Geschäftsmodelle*
M1	Wertangebote
M2	Kundensegmente
M3	Kanäle zu Partnern
M4	Kanäle zu Kunden
M5	Kommunikation zu Stakeholdern
M6	Beziehung zu Kunden
M7	Beziehung zu Partnern
M8	Schlüsselaktivitäten
M9	Schlüsselressourcen
M10	Schlüsselpartner
M11	Einnahmequellen

3.4 Anforderungen an plattformbasierte Geschäftsmodelle

Die Ergebnisse der Anforderungsgenerierung auf Basis des vorgestellten Vorgehensmodells mit zwei Unternehmen aus dem Bereich der Sharing Economy ist in nachfolgender Tabelle 3 dargestellt. Die erarbeiteten Anforderungen werden mit A1-9 bezeichnet.

Tabelle 3. Anforderungen an plattformbasierte Geschäftsmodelle

Nr.	*Anforderung*
A1	Vertrauen in Problemlösungskompetenz schaffen
A2	Attraktivität für Multiplikatoren gewährleisten
A3	Vertrauen in andere Stakeholder schaffen
A4	Asymmetrische Marktentwicklung gewährleisten
A5	Schutz der Intermediärfunktion der Plattform gewährleisten
A6	Positive Wahrnehmung der Stakeholder durch Teilhabe am Plattformökosystem erzeugen
A7	Zuverlässigkeit gewährleisten
A8	Verfügbarkeit geeigneter personeller und materieller Ressourcen gewährleisten
A9	Potenzial zur Generierung regelmäßiger und planbarer Einnahmen gewährleisten

Dieser Anforderungskatalog dient als Ausgangspunkt für das im Folgenden vorgestellte Vorgehensmodell zur Bewertung plattformbasierter Geschäftsmodelle (Forschungsfrage 3).

4 Vorgehensmodell zur Bewertung von Geschäftsmodellen

Auf Basis der erarbeiteten Anforderungen an plattformbasierte Geschäftsmodelle sollen ebensolche Geschäftsmodelle künftig auf mögliches Verbesserungspotenzial hin untersucht werden können. Abbildung 2 stellt den Ablauf eines solchen Workshops graphisch dar. Das Modell entspricht im Wesentlichen einer Umkehrung des in Kapitel 3.2 erläuterten Modells zur Generierung von Anforderungen und folgt grob dem Aufbau eines klassischen QFD Workshops. Nachfolgend werden die einzelnen Schritte näher erläutert.

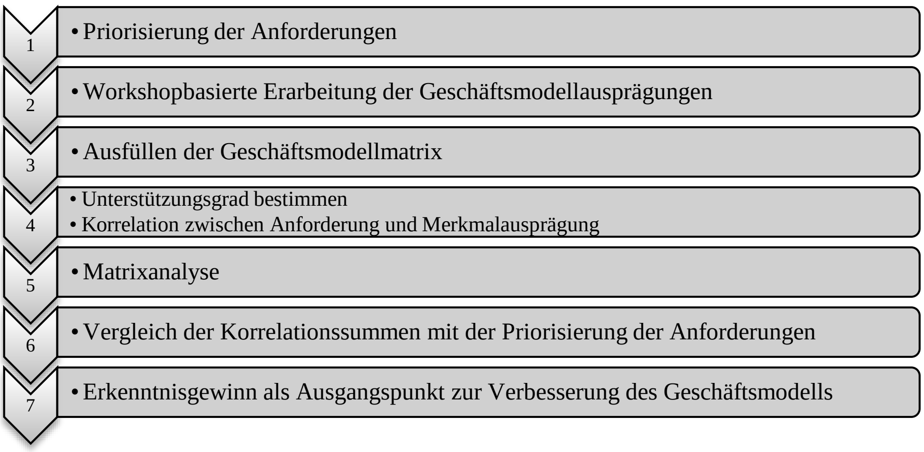


Abbildung 2. Vorgehensmodell zur Bewertung plattformbasierter Geschäftsmodelle

Da zu erwarten ist, dass sich die zu untersuchenden Geschäftsmodelle und das Unternehmensumfeld für das sie erarbeitet werden stark unterscheiden, ist entsprechend auch davon auszugehen, dass die Anforderungen an diese Geschäftsmodelle unterschiedlich bedeutend sein werden. Aus diesem Grund soll der Anforderungskatalog zunächst durch die Autoren des Geschäftsmodells priorisiert werden (Schritt 1). Dies kann einerseits dazu beitragen, dass sich die Workshopteilnehmer mit den Anforderungen auseinandersetzen, andererseits kann die Priorisierung in einem späteren Schritt zur Analyse der Geschäftsmodellmatrix genutzt werden. Im Anschluss sollen die Ausprägungen des Geschäftsmodells erarbeitet werden (Schritt 2). Dies kann workshopbasiert, beispielsweise mithilfe der Business Model Canvas oder den in Kapitel 3.3 vorgeschlagenen Geschäftsmodellmerkmalen geschehen. Die Ergebnisse sollen in Schritt 3 in die Geschäftsmodellmatrix eingetragen werden.

Nachfolgend soll der Unterstützungsgrad der Anforderungen durch die in Schritt 2 erarbeiteten Geschäftsmodellausprägungen diskutiert und mit den Werten 0, 3 und 9 bewertet werden (Schritt 4). In Schritt 5 soll die Matrix auf schwach gelöste Zeilen, also unzureichend gelöste Anforderungen, untersucht werden (siehe Kapitel 3.2). Mithilfe der in Schritt 1 erarbeiteten Priorisierungen der Anforderungen soll in Schritt 6 überprüft werden, ob die Merkmalsausprägungen des Geschäftsmodells dieser Priorisierung entsprechen. Dies kann Hinweise auf weiteren Änderungsbedarf aufzeigen. Im letzten Schritt (7) steht dann der Erkenntnisgewinn, der zur Überarbeitung und Verbesserung des Geschäftsmodells führen soll.

5 Kritische Würdigung

Mithilfe der vorliegenden Arbeit sollte eine Möglichkeit zur Bewertung plattformbasierter Geschäftsmodelle mit Quality Function Deployment erarbeitet werden. Hierzu wurden Mechanismen der Methode QFD genutzt, um Anforderungen an plattformbasierte Geschäftsmodelle zu erarbeiten, die künftig zu Bewertung solcher Geschäftsmodelle eingesetzt werden können. Diese Zielsetzung findet in der Formulierung der Forschungsfragen dieses Beitrags Niederschlag (Tabelle 1).

Bei der Beantwortung der ersten Forschungsfrage wurde evident, dass plattformbasierte Geschäftsmodelle alle Merkmale von klassischen Geschäftsmodellen aufweisen, jedoch darüber hinaus der großen Bedeutung der zusätzlichen Stakeholder Rechnung tragen müssen. Dies betrifft insbesondere das Management der Partner bzw. Komplementoren. Der Plattformanbieter sollte entsprechend nicht nur das eigene Geschäftsmodell im Blick haben, sondern auch sicherstellen, dass die auf seiner Plattform vertretenen Komplementoren selbst ein tragfähiges Geschäftsmodell betreiben können. Idealerweise enthält das Geschäftsmodell des Plattformanbieters bereits einen Blueprint für das Geschäftsmodell der Komplementoren.

Die identifizierten Geschäftsmodellmerkmale sind notwendig, um die spezifischen Anforderungen plattformbasierter Geschäftsmodelle lösen zu können. Diese Anforderungen betreffen insbesondere die Ausgestaltung der Beziehung zwischen den Stakeholdergruppen und die Attraktivität der dem Geschäftsmodell zugrundeliegenden Plattform für die relevanten Akteure.

Auf Basis dieser Ergebnisse konnte ein Vorgehensmodell für den Ablauf eines Bewertungsworkshops für plattformbasierte Geschäftsmodelle entwickelt werden.

Aufgrund von Restriktionen bezüglich der Bearbeitungszeit im Rahmen einer Abschlussarbeit sind Defizite bezüglich der Aussagekraft und Übertragbarkeit der erarbeiteten Ergebnisse zu erwarten. Eine erste Einschränkung ist der nicht beleuchtete Zusammenhang zwischen Geschäftsmodellmerkmalen. Dies kann die Analyse der Geschäftsmodellmatrix und damit die Formulierung genauerer Implikationen und Handlungsempfehlungen erschweren.

Weiterhin wurde bei der Wahl der verwendeten Geschäftsmodelleinteilung auf eine hohe praktische Durchführbarkeit Wert gelegt. Dies kann sich ebenfalls negativ auf den Detaillierungsgrad der erarbeiteten Ergebnisse auswirken. Weitere Defizite betreffen die Auswahl und Anzahl der teilnehmenden Unternehmen: Da es sich hierbei um Unternehmen in frühen Phasen ihrer Lebenszyklen handelt, ist der Erfahrungsschatz der am Workshop beteiligten Experten als eingeschränkt zu bewerten. Weiterhin war es aufgrund des hohen zeitlichen und personellen Aufwands der Anforderungsworkshops nicht möglich eine größere Anzahl an Unternehmen zu untersuchen. Entsprechend sind negative Auswirkungen auf die Allgemeingültigkeit der erarbeiteten Anforderungen zu erwarten. Ebenfalls aus Zeitgründen konnten die Korrelationen innerhalb der Matrix nicht in Zusammenarbeit mit den Experten erarbeitet werden, sondern wurden von den

Autoren dieser Arbeit auf Basis der Expertenaussagen vorgenommen. Die Beschränkung auf nur drei Korrelationswerte vereinfachte dies einerseits, führte aber ebenfalls zu Beeinträchtigungen hinsichtlich der Differenzierung der Ergebnisse.

Trotz dieser Defizite konnte die Eignung der Methode Quality Function Deployment für die Bewertung plattformbasierter Geschäftsmodelle gezeigt werden. Weiterhin betreffen die genannten Defizite hauptsächlich die operative Durchführung der Anforderungs- und Merkmalserhebung. Diese beeinträchtigen damit nicht die grundsätzliche Eignung des Verfahrens. Die positive Resonanz der teilnehmenden Experten lässt darüber hinaus auf das Potenzial dieses Verfahrens als Hilfestellung für junge Unternehmen schließen.

Die Ergebnisse der auf Basis des erarbeiteten Vorgehensmodells durchgeführten Workshops lassen auf sein Potenzial für weitere Untersuchungen schließen. Zur Minimierung der angesprochenen Defizite sollten weitere Workshops mit Unternehmen in reiferen Phasen durchgeführt werden. Die Ergebnisse können mit der Identifikation von möglichen Zusammenhängen der Geschäftsmodellkomponenten im *Dach* des House of Quality („WIE zu WIE“-Abhängigkeiten, vgl. Abbildung 1) weiter analysiert werden. Das Resultat sollte ein umfangreicherer und detaillierterer Anforderungskatalog sein. Zur Bildung eines solchen Kataloges können auch detailliertere Geschäftsmodellgliederungen wie bspw. die Service Business Model Canvas [23] genutzt werden.

Mithilfe dieser Anforderungen können Startups ihr Geschäftsmodell vor der Implementierung überprüfen. Die erarbeitete Bewertungsmethode kann somit potenziell einen Beitrag zur Vermeidung von gescheiterten Unternehmensgründungen leisten.

Das erarbeitete Verfahren bietet darüber hinaus auch Verwendungsmöglichkeiten über die betrachteten plattformbasierten Geschäftsmodelle hinaus: Unterstellt, dass diese eine spezifischere Teilmenge von Geschäftsmodellen insgesamt sind, könnten über die Anwendung des erwähnten Kano-Modells irrelevante Anforderungen identifiziert und aus der Betrachtung herausgenommen werden.

Literatur

1. Veit, D., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Spann, M., Kundisch, D., Leimeister, J.: Geschäftsmodelle – Eine Forschungsagenda für die Wirtschaftsinformatik. In: Wirtschaftsinformatik, vol. 56, pp. 55-64. Springer, Wiesbaden (2014)
2. Denning, S.: An Economy of Access is opening Business: Five Strategies for Success. In: Strategy & Leadership, vol. 4, pp. 14-21. Emerald, Bradford (2014)
3. Mautsch, L.M.: Softwareplattformen für Unternehmenssoftwareökosysteme. Eul, Lohmar (2015)
4. Jansen, S., Cusumano, M.: Defining Software Ecosystems: A Survey of Software Platforms and Business Network Governance. In: Jansen, S., Brinkkemper, S., Cusumano, M. (eds.) Software Ecosystems: Analyzing and Managing Business Networks in the Software Industry, pp. 13-28. Edward Elgar, Northhampton (2013)
5. Giardino, C., Wang, X., Abrahamsson, P.: Why early-stage Software Startups fail: A behavioral Framework. In: Proceedings of the 5th International Conference on Software Business. Paphos (2014)
6. CBInsights: The Top 20 Reasons Startups fail, http://www.cbinsights.com/research-reports/The-20-Reasons-Startups-Fail.pdf (Accessed: 16.08.2016)

7. Teece, D.J.: Explicating Dynamic Capabilities: The Nature and Microfoundations of (sustainable) Enterprise Performance. In: Strategic Management Journal, vol. 28, pp. 1319-1350. Wiley, Malden (2007)
8. Cusumano, M.A.: Staying Power – Six enduring Principles for managing Strategy and Innovation in an uncertain World (Lessons from Microsoft Apple Intel Google Toyota and more). University Press, Oxford (2010)
9. Burkhart, T., Krumeich, J., Werth, D., Loos, P.: Analyzing the Business Model Concept – A comprehensive Classification of Literature. In: Proceedings of the 32nd International Conference on Information Systems. Shanghai (2011)
10. Baden-Fuller, C., Morgan, S.: Business Models as Models. In: Long Range Planning, vol. 43, pp. 156-171. Elsevier, Oxford (2010)
11. Rusnjak, A.: Entrepreneurial Business Modeling im Kontext einer erfolgsfaktorenorientierten Strategie-Entwicklung – Entwicklung eines Vorgehensmodells, Frameworks und Werkzeugs zur semiformalen Modellierung und Visualisierung früher Anforderungen von der Idee bis zum Produkt bzw. Startup im e/MCommerce. Dissertation an der Universität Leipzig. Leipzig 2012
12. Osterwalder, A.: The Business Model Ontology. Dissertation an der Universität Lausanne. Lausanne (2004)
13. Al-Debei, M.M., El-Haddadeh, R., Avison, D.: Defining the Business Model in the New World of Digital Business. In: Proceedings of the Fourteenth Americas Conference on Information Systems. Association for Information Systems (AIS), Toronto (2008)
14. Gawer, A.: Bridging differing Perspectives on technological Platforms: Toward an integrative Framework. In: Research policy, vol. 43, pp. 1239-1249. Elsevier, Oxford (2014)
15. Rochet, J.-C., Tirole, J.: Platform Competition in two-sided Markets. In: Journal of the European Economic Association, vol. 4, pp. 990-1029. Wiley Blackwell, Hoboken (2003)
16. Dillerup, R., Stoi, R.: Unternehmensführung. Vahlen, München (2013)
17. Herzwurm, G., Schockert, S., Mellis, W.: Qualitätssoftware durch Kundenorientierung – Die Methode Quality Function Deployment (QFD): Grundlagen, Praxis und SAP R/3 Fallbeispiel. Vieweg, Braunschweig (1997)
18. Cohen, L.: Quality Function Deployment – How to make QFD work for you. Addison-Wesley, Reading (1995)
19. Puschmann, T., Alt, R.: Sharing Economy. In: Business & Information Systems Engineering, vol. 58, pp. 93-99. Springer, Wiesbaden (2015)
20. Osterwalder, A., Pigneur, Y.: Business Model Generation – Ein Handbuch für Spieleveränderer und Herausforderer. Campus (2011)
21. Strahinger, S. (eds.): Geschäftsmodelle der IT-Industrie. Dpunkt, Heidelberg (2013)
22. Saatweber, J.: Kundenorientierung durch Quality Function Deployment: Produkte und Dienstleistungen mit QFD systematisch entwickeln. Symposion, Düsseldorf (2011)
23. Zolnowski, A., Weiß, C., Böhmann, T.: Representing Service Business Models with the Service Business Model Canvas – The Case of a Mobile Payment Service in the Retail Industry. In: Proceedings of the Forty-Seventh Annual Hawaii International Conference on System Sciences. Waikoloa (2014)
24. Akao, Y., Mazur, G.H.: The leading Edge in QFD: Past, Present and Future. In: International Journal of Quality & Reliability Management, vol. 20, pp. 22-35 (2003)
25. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. In: MIS Quarterly, vol. 28, pp. 75-105 (2004)
26. Gregor, S., Hevner, A.R.: Positioning and presenting Design Science Research for maximum Impact. In: MIS Quarterly, vol. 37, pp. 337-355 (2013)

A Picture is Worth a Thousand Words: Visual Model Evaluation in Data Science Applications

Dennis Eilers[1], and Michael H. Breitner[1]

[1] Leibniz Universität Hannover, Information Systems Institute, Hannover, Germany
{eilers,breitner}@iwi.uni-hannover.de

Abstract. Besides programming and mathematical or statistical skills, domain knowledge about the investigated problem is an important factor in data science. In practical applications, however, there often exist a gap between data scientists who have the technical skills in advanced analytics methods and the domain experts like managers and decision makers with substantial knowledge in their field. Less participation of domain experts during the complex model building and evaluation process of the data science pipeline can lead to acceptance problems and prejudices against the results. Moreover, data scientists depend on the domain knowledge of experts when it comes to evaluation and problem identification of their models. In this paper we address this issue by introducing an easy-to-understand heat map visualization technique for model evaluation. It enables all parties to discuss the analysis on the same level of complexity. The benefits are demonstrated based on a real world business example.

Keywords: Data Science, Visualization, Machine Learning Model Evaluation, Heat Maps

1 Introduction and Motivation

Data science is often described as a combination of mathematics/statistics, programming and domain knowledge [1]. In a business context, data analysis is the task of specialized professionals who have substantial knowledge about the data bases, analysis tools and mathematical/statistical models. It is widely discussed in the literature how Decision Support Systems (DSS) based on data analytics can improve management performance [2-3]. From the opposite perspective, regarding the definition of data science, it is also an important question how to make use of managers' domain knowledge during the whole analysis process. Therefore, communication between the different expert groups is essential for team performance [4]. Data visualization can be a powerful tool for this purpose. There exist many studies about visualization for data exploration [5-6], result communication [7-8] and decision support [9-10]. Using visualization as a communication instrument between data scientists and domain experts during the actual model building and evaluation steps in the analysis pipeline is unexplored so far. The lack of participation in this technical and complex but crucial phase can lead to prejudices and acceptance problems on the part of managers about the final results. Especially the use of advanced analytics techniques

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Eilers, D.; Breitner, M.H. (2017): A Picture is Worth a Thousand Words: Visual Model Evaluation in Data Science Applications, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1489-1500

like Artificial Neural Networks (ANN) then often lead to the common “black-box” criticism. Using the experience of managers even with lesser technical skills can also help the data scientists during the model building and especially during the model evaluation phase. Problems with the model performance which are not attributable to the methods or algorithms (e.g. an omitted variable bias) can be identified much easier with the help of expert knowledge, which leads to better results in an iterative process of model building, evaluation and adjustment. Therefore, we are interested in the question of how visualization can help to integrate managers in the model building and evaluation process of the data science pipeline to generate acceptance, confidence and trust [11-12] and also to make better use of domain knowledge for model improvements.

This paper is motivated by a real world application in cooperation with a large German car manufacturer. The data analysis problem is about improving the residual value forecast for used cars to support the decision makers in the leasing business of the company. Based on discussions with the data scientists and domain experts (managers/decision makers) we propose a new visualization technique for machine learning model evaluation using the familiar idea of heat maps to address the aforementioned question. The contribution can be summarized as follows.

- From a manager perspective: The proposed heat map visualization for model evaluation provides a familiar and easy interpretable way to assess the performance of the analysis. Experience and domain knowledge can be incorporated during the discussions about the analysis with the data scientists. The participation in this phase can help to develop acceptance and trust in advanced analytics methods which lead to more confident decision making.
- From a data scientist perspective: The intuitive visualization helps to make better use of the domain knowledge from experts to identify the problems of the models. It also helps to identify the exact regions of bad performance in the data. The method is applicable for any model complexity and also possible to use with Big Data sets.

The next section provides an overview about the concrete problem statement in the real world business application. Section 3 describes the solution approach and the implementation of the proposed heat map model evaluation technique. In section 4, the results and limitations of this approach are discussed. Sections 5 concludes and provides an outlook for further research.

2 Set the Stage

This section describes the data analysis task in the real world business application and illustrates the problem statement. The general question of this analysis project is how residual values of used cars evolve over time [13-14]. This question is crucial in the leasing business, because the leasing rate (the amount of money a customer has to pay during the leasing period) is determined by the estimation of the future residual value of a car. The loss in value over time needs to be compensated by this rate. An overestimation of the residual value leads to a lower leasing rate (competitive advantage) but also generates a loss when reselling the car after the leasing period. A conservative underestimation leads to high leasing rates and therefore generates an avoidable competitive disadvantage. With more than 150.000 recorded leasing contracts and resales over four years (2011-2014) from a major German car manufacturer, this is a typical regression problem where machine learning models can be used to explain the residual values based on features of the car like the age, the mileage or the fuel type. The trained models can be used for future contracts to determine a more accurate estimation of the expected residual value. The results affect different stakeholders, namely the car manufacturer, the leasing bank, the authorized dealers and the customers.

In such a practical application the experience and domain knowledge from managers and decisions makers about the leasing market and its special characteristics are an important factor in the whole analysis process. The goal is to achieve a closer cooperation between the data science experts and the many different experts who have in depth knowledge and experience in the field of leasing contracts and the leasing market. For this forecasting application we use ANNs which are suitable for noisy and nonlinear data [15-16]. The common criticism of ANNs is their "black-box" characteristic. After training the ANNs with data from the past (not the focus of this paper) the models are able to produce a forecast of previously unseen data (new contracts). ANNs in general and the training process in particular are highly complex mathematical optimization problems which are difficult to understand for professionals with fewer technical skills. General rejection and prejudices can be the result. But in the end, the decisions are made by humans who have to accept and trust the forecasts. On the other hand, the data scientists face the problem of incorporating the domain experts during the analysis to get the necessary domain knowledge about the data and reveal possible problems with the models which can only be identified by experts who have probably less mathematical and statistical skills (e.g. identifying omitted variables).

The general idea is to use heat maps, a familiar tool for most business people [17], to incorporate all different kinds of experts in the evaluation phase of the initial analysis, which supports iterative adjustments and improvements. The goal is to provide a method where domain experts with less technical skills, managers and data scientists can discuss the performance of the current approach on the same level of complexity. Problems with the analysis like missing data, omitted input variables or nonlinearities can be identified together, whereby all participants are actively involved in the process of deriving new or improved models. The next section describes the implementation of the proposed heat map method for model evaluation and illustrates its functionality.

3 Visual Model Evaluation with Heat Maps

Heat maps are widely used in research fields like geology or meteorology [18] and play an increasingly important role in business applications. Buehler and Pritsch use heat maps to present results of risk management models in a more intuitive and comprehensible way [17]. By providing easy to understand heat map visualizations of different risk categories and business units they claim to make risk more transparent. Heatmaps could thus contribute to a dialogue between the board of directors, senior management and business unit leaders. Köpp et al. propose a heat map visualization technique for applications where forecast distributions of several future time steps are generated [19]. For an easier interpretation, forecast ensembles are often aggregated by the mean or median, which reduces the information to a single forecast line. The proposed heat map intuitively visualizes areas in the ensemble with high and low activity, which represents the uncertainty of the models and therefore facilitates decision making (for example the identification of the best point in time to buy a certain commodity). This is an example of how heat maps can be used to aggregate data without losing important information.

In this section, we propose a new technique for model evaluation using heat map visualizations of forecasting errors/residuals. In most data science oriented studies, model evaluation is based on some common performance measures like the Root Mean Square Error (RMSE) or the Mean Absolute Error (MAE) [20]. These performance measurements reduce the information about the model quality to a single number. These measures are usually used to compare different models or model specifications with each other in order to choose the best performing one. This common approach has the drawback that the information where exactly the model performs good or bad in the input and output space, cannot be observed. In addition, these rather technical evaluations of the analysis have no value for incorporating domain knowledge in the evaluation and are unable to generate trust in results of probably highly complex "black-box" methods.

Using heat maps for residual visualization is a new approach and has several advantages compared to conventional methods. With heat maps it is possible to incorporate a further dimension, because the information about the residuals can be visualized by a color scale. Therefore, it is possible to use a two-dimensional coordinate system with two variables of the input or output space on the axes. These "cuts" through the data enable the user to visualize the residuals in the same data space as a conventional scatterplot or density plot of the data points. This can help to detect hidden anomalies in the data/model when comparing the residual heat maps with scatterplots or density plots which are in the same range and size. These hidden anomalies can be, for example, a bad performance at the borders of the distribution or a special region like a data gap in the input space, which leads to bad performance in the results. This can give hints for an omitted variable bias or an insufficiently incorporated, nonlinear relationship between variables. The representation of the error terms in the form of a heat map remains clear and comprehensible even with a large number of data points, since the general appearance is independent of the number of data used to generate the visualization. Therefore, the method is also suitable for Big Data applications. However, in the case of large data sets, data aggregation takes place but without losing information about the performance in different regions of the data space, as is the case with single performance measures.

To illustrate the technique, let's consider 500 samples from a two-dimensional standard normal distribution. This stylized example serves as an artificial use case with known distribution and properties of the data points to validate the functionality of the method within a controlled environment. The two dimensions (x and y) represent two continuous input variables (features) of a machine learning regression model. This information can be used to generate a scatterplot of the data points to get an idea about the distribution and possible correlations. Figure 1 shows a smoothed scatterplot of this example.

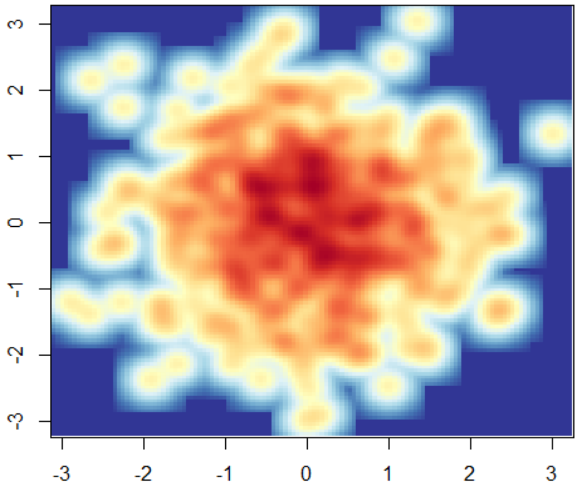


Figure 1. Smoothed scatterplot of 500 random samples of a two-dimensional standard normal distribution.

The residuals which correspond to these data points are now visualized in the exact same data space using a color scale. The assumption is, that each residual is representative for a specified region around its position. For example, it might be the case that a forecasting model predicts the realized values very well in one region of the data space (residuals in that area are rather low) and in another region the forecasting accuracy is less accurate (residuals in that area are rather high). Since the position of the data points are typically not on a regular grid, a method for weighting and smoothing is necessary. By using a kernel around the data points, each residual gets a specific range of influence. At the position of the data point, the influence of the residual should be high, while it decreases with increasing distance. The goal of using the kernel approach is to generate a regular grid with the weighted influence of the residuals on each grid point. A resulting discrete and regular grid is defined as a matrix of the dimension (m, n) within the same variable space. There are several other interpolation methods for irregular grids like kriging and inverse distance weighting [21], but in this use case, an actual weighting of the residuals has to be performed which must also work with data points at the exact same position (high density areas in the data). Each residual has a defined location in the variable space. In our approach, we use a two-dimensional Gaussian kernel with

$$\mu = \binom{\mu_A}{\mu_B} \text{ and } \Sigma = \begin{pmatrix} \sigma_A^2 & \rho\sigma_A\sigma_B \\ \rho\sigma_A\sigma_B & \sigma_B^2 \end{pmatrix} \quad (1)$$

where $\begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix} := \begin{pmatrix} x_i \\ y_i \end{pmatrix}$ for each new residual i. The covariance matrix Σ can be set by the user but for the default value we specify equation 2 with size s = 1000 as control parameter for the expansion of the kernel and zero correlation

$$\Sigma_{default} = \begin{pmatrix} \frac{(\max(X)-\min(X))^2}{s} & 0 \\ 0 & \frac{(\max(Y)-\min(Y))^2}{s} \end{pmatrix} \tag{2}$$

The density function in the special two-dimensional case, for a specific residual i, at the grid points (m, n), can be defined as

$$f^i(m,n) = \frac{1}{2\pi\sigma_A\sigma_B\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{(n-\mu_A)^2}{\sigma_A^2} + \frac{(m-\mu_B)^2}{\sigma_B^2} - \frac{2\rho(n-\mu_A)(m-\mu_B)}{\sigma_A\sigma_B}\right)\right) \tag{3}$$

For each residual, the position (x_i, y_i) is defined as the mean of the density function and each grid point is inserted into the generated function 3. This results in a list of weights (according to the two-dimensional normal density) for each residual, for each grid point. The threshold parameter t can be set to specify the minimal value of a weight, otherwise this residual has no influence on this specific grid point. The weight $w_{m,n}^i$ for residual i at grid point (m, n) is defined by

$$w_{m,n}^i = \begin{cases} f^i(m,n), & \text{if } f^i(m,n) \geq t \\ 0, & otherwise \end{cases} \tag{4}$$

If the distance between the residual and the grid point is too high (grid point inserted into the normal density kernel, with mean equals the position of the residual, results in a value below the threshold) the influence for this specific value is set to zero. Figure 2 illustrates the weighting procedure.

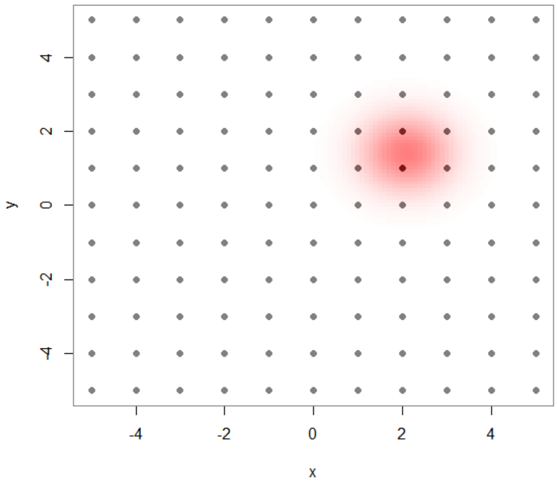


Figure 2. An example of one residual inserted into the regular grid of the heat map. The position of the residual in the data space is equal to the mean of the two-dimensional normal density.

The actual weighting for each residual r_i at each grid point is performed according to equation 5.

$$g_{m,n} = \begin{cases} \sum_{i=1} \left(r_i \frac{w_{m,n}^i}{\sum_{i=1} w_{m,n}^i} \right), & \text{if } \sum_{i=1} [w_{m,n}^i > 0] \geq c \\ NA, & \text{otherwise} \end{cases} \quad (5)$$

The cutting parameter c can be set to allow only those grid points to be colored that have at least a specified number of $w_{m,n}^i \geq c$ which means that more than one residual has an influence on that specific grid point. This can help to make the result more robust, because single value outliers are mitigated, but at the expense of greater information loss in low-density areas. After assigning a value $g_{m,n}$ to each point in the regular grid, the heat map can be produced. To illustrate the resulting heat maps, a third sample of 500 standard normal distributed values represents the artificial residuals corresponding to each (x, y) point in the data space. Figure 3 shows the result of this stylized example.

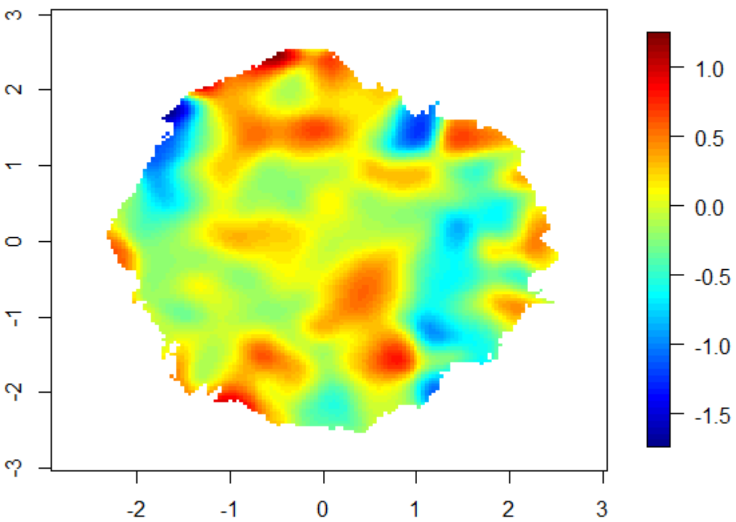


Figure 3. The figure shows a residual heat map which is generated according to the described procedure. In this case, the residuals are normally distributed which results in a heat map with no clear patters.

The red regions represent the high residual areas, while in the blue regions the residuals are rather small. In this example, there is no clear pattern apparent because the residuals are indeed normally distributed over the whole data space. This should be the case if the (forecasting) models work well in each area of the data space.

To illustrate a case when a specific region is biased, the residuals are now artificially manipulated by increasing the residuals located within the range of -0.1 and 0.1 on the x-axis according to equation 6.

$$r_i = \begin{cases} r_i + 2, & \text{if } -0.1 \le x_i \le 0.1 \\ r_i, & \text{otherwise} \end{cases} \quad (6)$$

Figure 4 shows the resulting heat map with the artificially biased region. A clear pattern is now visible, indicated by the red region. Such patterns would be expected if the model performs bad, especially in a particular region, for example because of an omitted variable or a nonlinearity which was not incorporated.

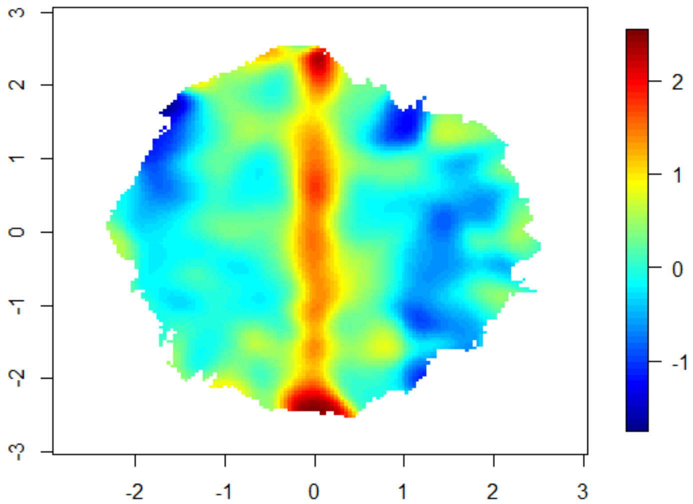


Figure 4. The residuals are now artificially manipulated by increasing the residuals located within the range of -0.1 and 0.1 on the x-axis. A clear pattern is visual now indicated by the red region. Such patterns would be expected if the model performs bad in a particular region for example because of an omitted variable or a nonlinearity which was not incorporated.

4 Discussion and Limitations

The method described in the previous section is now applied to the real world problem of forecasting the resale price of used cars. This example only serves as illustration purpose how visualization can be used in a business application and help to reveal previously unseen problems. After training the ANN models with the available input data, we visualize the resulting model errors on a validation dataset with 4500 hold out samples which were not represented in the training process. First, figure 5 presents a smoothed scatterplot of the observed residuals in the data space with the two variables “sale price” and “age” of the cars. The variable “age” is scaled and centered around zero (best practice for input variables of ANNs [22]), while the variable “sale price” is defined as the ratio of the initial list price and achieved market value. Figure 6 shows the resulting residual heat map (residuals as absolute values) of the same data set.

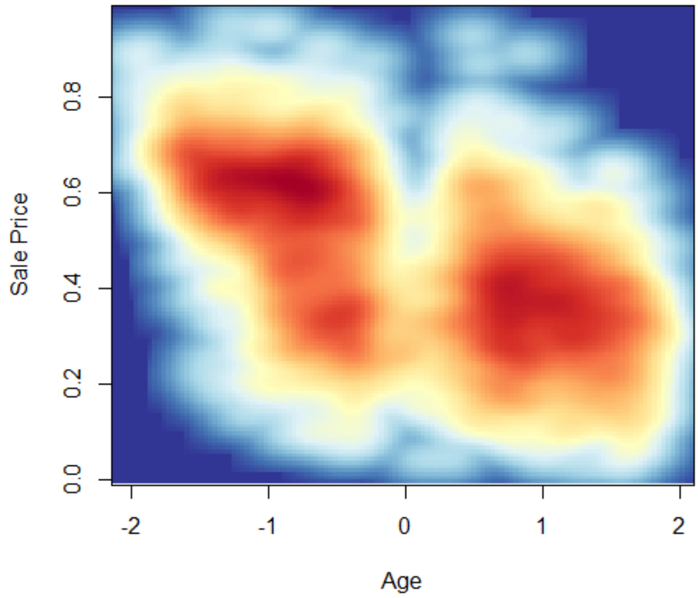


Figure 5. Smoothed scatterplot of the observed residuals in the data space with two variables: the sale price and age of the cars.

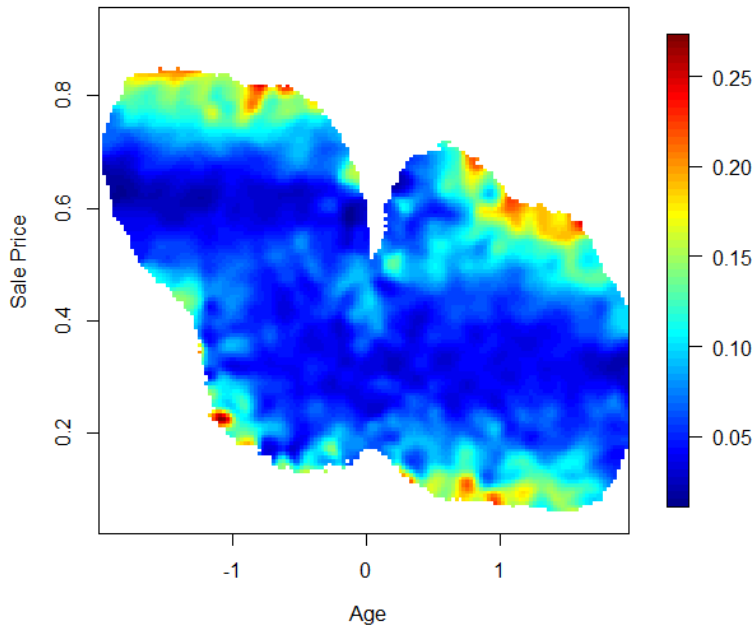


Figure 6. A systematic pattern in the residuals can be observed. In the upper and lower regions of the variable “sale price”, the residuals tend to be higher than in in the more centered parts.

A systematic pattern in the residuals of the ANN model can be observed. In the upper and lower regions of the variable “sale price” the residuals tend to be higher than in in the more centered parts. This can be a hint about information which are not incorporated in the model. Discussions with experts and managers about the observed results in this visualization reveal a plausible explanation at least for the upper regions of the sale price. In the data, no enhanced acquisition costs are reported. For example, a crane which was installed later on a used car increases the value of the whole car significantly, even if the residual value of the car decreases normally. When such cars are resold, no information about this extra equipment will be reported and the residual value of this car seems to be unusually high. The respective data points which would otherwise be marked as simple outliers can now be identified as patterns which are subject to a systematic bias induced by an omitted explanatory variable. Future data collection will incorporate information about enhanced acquisition costs.

From a manager perspective, this visualization technique enables the participation in the complex process of model building, evaluation and adjustment, even with lesser mathematical and statistical skills. They can assess the performance of the models within a familiar environment and actively discuss practice-oriented ideas about the problems of the advanced analytics models with the data science experts on the same complexity level. The participation in the “black-box” phase is also the basis for trust and acceptance of the results which can lead to more confident decision making.

From a data scientist perspective, this method can be used to uncover hidden problems in the data or algorithms based on the domain knowledge of experts in the field. This approach also offers a new way of visualizing the model error within the same data space as scatterplots or density plots. For example, comparing density plots and heat maps can provide a first insight about the performance in regions with less data or data gaps. This can give a hint if the problem is due to the data (possible omitted variables) or the used method (a nonlinearity which is not correctly incorporated). It provides much more valuable information than single performance measures like the RMSE. In contrast to conventional residual analytics methods like simple residual plots, the heat map can be used with an unlimited number of data points (applicable for Big Data machine learning models) in the evaluation. Residual plots are hardly interpretable if the number of data points increases. The color scale and the weighting function of the heat map allows to visualize many patterns without losing information about the model performance in different regions of the data space.

Like any approach, the heat map visualization faces some difficulties and drawbacks. Big Data machine learning applications often contain a large number of input factors, of which only two can be plotted simultaneously in one coordinate system. Scatterplot matrices can be used to represent all variables pairwise, but this also leads to a more complex visualization as the number of factors increases. In the described example, also only continuous variables are investigated within a regression problem. The information gain for one or even two discrete or binary variables on the axes is questionable. The method is also limited to regression problems. Classification problems can be incorporated by some adjustments on the heat map color scale (binary representation with two colors for true or false) for future tests. As mentioned in section 3, several parameters have to be set properly for achieving good results. These parameters could have a tremendous influence on the heat map appearance. It is not a plug and play solution and users need substantial knowledge about the backend of the system. It is therefore once again the responsibility of the data scientists to generate meaningful presentations which can also lead to skepticism. Wrong parameter settings could have tremendous negative effects regarding trust and decision making performance. The advantage that decisions can be made with more confidence using visualizations can also be a drawback if these approaches lead to overconfidence in the models. The principle "trust but verify" is important with any model, but the visualizations do not help to understand the mathematical structure or the training process in general. An important problem is how to measure the actual benefits and the role of trust, acceptance and confidence [11-12] when using such visualizations in the analysis process. In a real world application, hardly any quantifiable indicator exists for the performance improvements induced by this visualization except a qualitative argumentation.

5 Conclusion

In this paper, a heat map visualization for model evaluation is introduced. The overall goal is a better integration of managers and decision makers in the model building and evaluation process of the data science pipeline in order to generate trust and acceptance and furthermore, to make use of the experts' domain knowledge in this phase. Based on a real world business example, the benefits of this method were discussed. The visualization technique allows domain experts to actively participate in the technical and complex model building and evaluation process which helps to reveal a systematic bias induced by an omitted explanatory variable. Data scientists are equipped with a tool for visualizing model errors in the same data space as two dimensional scatterplots to identify the exact regions where the models perform good or bad. The method is applicable even for Big Data machine learning models.

In further research we are interested in better understanding the actual benefits of visualization for data science in general. One approach is the laboratory experiment, in order to measure the results of an analysis process in a controlled environment by providing different visualization tools for control and treatment groups. This also allows to measure the differences in trust, acceptance and confidence of the subjects. In general, this paper aims to encourage the use of visualization in data science and hopefully initiates more research in this important area.

References

1. Conway, D.: The Data Science Venn Diagram, http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram (Accessed: 25.09.2016)
2. Shim, J.P., Warkentin, M., Courtney, J.F., Power, D.J., Sharda, R., Carlsson, C.: Past, present, and future of decision support technology. Decision Support Systems. 33, 111–126 (2002)

3. Power, D.J.: Decision Support Systems: A Historical Overview. In: Handbook on Decision Support Systems 1. pp. 121–140. Springer Berlin Heidelberg (2008)
4. Bresciani, S., Eppler, M.J.: The Benefits of Synchronous Collaborative Information Visualization: Evidence from an Experimental Evaluation. IEEE Transactions on Visualization and Computer Graphics. 15, 1073–1080 (2009)
5. Keim, D.A.: Information visualization and visual data mining. IEEE Transactions on Visualization and Computer Graphics. 8, 1–8 (2002)
6. Keim, D.A., Panse, C., Sips, M., North, S.C.: Visual data mining in large geospatial point sets. IEEE Computer Graphics and Applications. 24, 36–44 (2004)
7. Kelleher, C., Wagener, T.: Ten guidelines for effective data visualization in scientific publications. Environmental Modelling & Software. 26, 822–827 (2011)
8. Sun, G.-D., Wu, Y.-C., Liang, R.-H., Liu, S.-X.: A Survey of Visual Analytics Techniques and Applications: State-of-the-Art Research and Future Challenges. J. Comput. Sci. Technol. 28, 852–867 (2013)
9. Al-Kassab, J., Ouertani, Z.M., Schiuma, G., Neely, A.: Information visualization to support management decisions. Int. J. Info. Tech. Dec. Mak. 13, 407–428 (2014)
10. Franz, M., Scholz, M., Hinz, O.: 2D versus 3D Visualizations in Decision Support – The Impact of Decision Makers' Perceptions. ICIS 2015 Proceedings. (2015)
11. Becker, J., Heddier, M., Öksüz, A., Knackstedt, R.: The Effect of Providing Visualizations in Privacy Policies on Trust in Data Privacy and Security. In: 2014 47th Hawaii International Conference on System Sciences. pp. 3224–3233 (2014)
12. Sacha, D., Senaratne, H., Kwon, B.C., Ellis, G., Keim, D.A.: The Role of Uncertainty, Awareness, and Trust in Visual Analytics. IEEE Transactions on Visualization and Computer Graphics. 22, 240–249 (2016)
13. Wu, J.-D., Hsu, C.-C., Chen, H.-C.: An expert system of price forecasting for used cars using adaptive neuro-fuzzy inference. Expert Systems with Applications. 36, 7809–7817 (2009)
14. Lessmann, S., Listiani, M., Voß, S.: Decision Support in Car Leasing: A Forecasting Model for Residual Value Estimation. ICIS 2010 Proceedings. (2010)
15. Bishop, C.M.: Neural networks for pattern recognition. Oxford university press (1995)
16. Schocken, S., Ariav, G.: Neural networks for decision support:: Problems and opportunities. Decision Support Systems. 11, 393–414 (1994)
17. Buehler, K.S., Pritsch, G.: Running with risk. McKinsey Quarterly. 40–49 (2003)
18. Hagh-Shenas, H., Kim, S., Interrante, V., Healey, C.: Weaving Versus Blending: a quantitative assessment of the information carrying capacities of two alternative methods for conveying multivariate data with color. IEEE Transactions on Visualization and Computer Graphics. 13, 1270–1277 (2007)
19. Köpp, C., Mettenheim, H.-J., Breitner, M.H.: Decision Analytics with Heatmap Visualization for Multi-step Ensemble Data - An Application of Uncertainty Modeling to Historical Consistent Neural Network and Other Forecasts. Business & Information Systems Engineering. 6, 131–140 (2014)
20. Witten, I.H., Frank, E., Hall, M.A.: Data Mining: Practical Machine Learning Tools and Techniques, Third Edition. Morgan Kaufmann, Burlington, MA (2011)
21. JIN, G., LIU, Y., NIU, W.: Comparison between Inverse Distance Weighting Method and Kriging [J]. Journal of Changchun University of Technology. 3, (2003)
22. LeCun, Y.A., Bottou, L., Orr, G.B., Müller, K.-R.: Efficient BackProp. In: Montavon, G., Orr, G.B., and Müller, K.-R. (eds.) Neural Networks: Tricks of the Trade. pp. 9–48. Springer Berlin Heidelberg (2012)

Die Rolle von Moral Disengagement in Crowdlending

David Schwarz, Philipp Haas

University of St.Gallen, Institute of Information Management, St. Gallen, Switzerland
{david.schwarz.home}@gmail.com
{philipp.haas}@unisg.ch

Abstract:
In den letzten Jahren konnte sich Crowdlending als neuartige Form der Kreditvergabe zunehmend etablieren. Dabei hat es sich mit 11,2 Milliarden US-Dollar Kreditvolumen in 2014 zur bedeutendsten Form der Finanzierung innerhalb des Crowdfunding Paradigmas entwickelt und sammelt dabei mehr Investitionsmittel ein als alle anderen Formen des Crowdfundings zusammen. In zunehmendem Masse wird Schwarmfinanzierung damit relevant in Wirtschaft und Politik. Gleichzeitig hat sich Crowdfunding und Crowdlending als hochrelevantes Forschungsthema in der Wirtschaftsinformatik etabliert. Während sind die bisherige Forschung dabei weitestgehend auf den Einfluss sozialer Netzwerke, der Bedeutung von sozialem Kapital sowie subjektivem Risiko und Antiselektion konzentriert, ist der Einfluss von moralisch-ethischen Steuerungsmechanismen bisher noch nicht weiter untersucht. Anhand einer Mediationsanalyse der Daten einer Umfrage mit 118 Teilnehmern konnte explorativ-quantitativ bewiesen werden, dass Moral Disengagement Mechanismen die Tendenz, unmoralische Investitionsentscheidungen zu treffen, beeinflusst. Die Studie legt damit die Grundlagen für weitere Forschung im Bereich des Crowdlending, um besser zu verstehen, wie Crowdlending Plattformen gestaltet werden müssen, um unethische und unmoralische Investitionen zu verhindern.

Keywords: Crowdfunding, Crowdlending, Peer-to-Peer Lending, Moral Disengagement

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Schwarz, D.; Haas, P. (2017): Die Rolle von Moral Disengagement in Crowdlending, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1501-1512

1 Einführung

In den letzten Jahren hat das Thema Crowdlending aufgrund hoher Wachstumsraten, neuen Erkenntnissen in der Forschung und zunehmender Berichterstattung in den Medien an Bedeutung gewonnen. Auf Crowdlending Plattformen können Kapitalnehmer ein Kreditgesuch einstellen, für die der potentielle Kreditgeber einen festgelegten Zinssatz als Risikoprämie erhält. So wurden über die weltweit führende Crowdlending Plattform *Lending Club* bereits Kredite mit einem Volumen von über 11,2 Milliarden US-Dollar vergeben mit stark steigender Tendenz [1]. Crowdlending gewinnt in der Wirtschaftsinformatik-Forschung immer mehr an Bedeutung, da erst das Internet und die Möglichkeiten des Web 2.0 den Erfolg der Crowdlending-Plattformen ermöglichen.

Kreditgeber bevorzugen im Crowdlending überwiegend Investitionen mit hohem Zinssatz und zeichnen sich durch überdurchschnittliche Risikobereitschaft aus [2]. Durch den einfachen Zugang zu Kapital auch für risikoreichere Zielgruppen und die attraktiven Renditen fördert Crowdlending zum einen die Gefahr der Überschuldung der Kreditsuchenden und zum anderen unmoralisches Investitionsverhaltens der Kapitalgeber [3]. Zum Beispiel werden im Crowdlending Kredite oft für die Rückzahlung von anderen Krediten wie beispielweise Kreditkartenschulden verwendet [4]. Unter unmoralischem Investitionsverhalten versteht man in diesem Kontext die Kreditvergabe auf Crowdlending Plattformen zum Nachteil des Kreditsuchenden. Dies kann sich in marktunüblichen und unvorteilhaften Kondition wie überzogenen Zinssätzen ausdrücken oder der Vergabe von Krediten trotz offensichtlich drohender Überschuldung. Hierbei umgeht der Kreditgeber die eigenen moralischen Selbstregulierungsmechanismen, was durch die subjektiv wahrgenommene Anonymität auf einer Internetplattform gefördert wird.

Aus diesem Grund wandelt sich die zuerst positive öffentliche Wahrnehmung von Crowdlending, was zu zunehmender regulatorischer Aufsicht führt. So limitiert beispielsweise das deutsche Bürgerliche Gesetzbuch den möglichen Zinssatz. Sittenwidrig ist dabei eine Kreditvergabe, wenn der effektive Jahreszins den Vergleichszins relativ um mehr als 91 Prozent überschreitet [5] und eine Notlage oder Unerfahrenheit des Darlehensnehmers ausgenutzt wird. Wenngleich die Gesetzgebung in Deutschland unethisches Investitionsverhalten vergleichsweise effektiv unterbinden kann, so ist die Regulierung in anderen Ländern nicht so streng.

Es ist somit von grosser Bedeutung zu verstehen, wie psychologische Mechanismen es ermöglichen, unmoralische Investitionen zu rechtfertigen. Zu diesem Zweck wird ein Strukturmodel anhand der Moral Disengagement Theorie nach Bandura et al. [6, 7] entwickelt, in dem Moral Disengagement als Mediator die Beziehung zwischen Moralischer Identität beziehungsweise Motive für finanziellen Gewinn und unmoralischem Investitionsverhalten moderiert. Zwar wurde in den letzten Jahren zum Thema Investorenverhalten bei Crowdlending geforscht und publiziert, allerdings konzertiert sich die Forschung dabei auf den Einfluss sozialer Netzwerke [8], subjektives Risiko (Moral Hazard) und Adverser Selektion [9]. Zur Frage des Einflusses moralischer Bedenken und der Überwindung derselben existiert bisher nur unzureichende Forschung.

Aus diesem Grund untersuchen wir in dieser Arbeit, wie Moral Disengagement im Crowdlending begründet wird und beschreiben, wie Individuen das persönliche moralische Selbstregulationssystem umgehen, welches unmoralisches Investitionsverhalten eigentlich reguliert. Dies öffnet eine neue Perspektive nicht nur in der Crowdlending Forschung, sondern lässt sich auch auf die Crowdfunding Forschung allgemein ausweiten. Hierfür wurde eine Befragung unter Kapitalgebern durchgeführt und mittels Mediationsanalyse ausgewertet. Die Ergebnisse zeigen, dass Investoren Moral Disengagement einsetzen, um das moralische Selbstregulationssystem zu überwinden und unmoralische Investition zum Ziele der Gewinnmaximierung zu tätigen.

Nach einem Überblick über die theoretischen Hintergründe zu Moral Disengagement und Crowdfunding stellen wir unser methodisches Vorgehen vor. Anschliessend werden die Ergebnisse der Mediationsanalyse analysiert und die Hypothesen ausgewertet. Schlussendlich werden theoretische und praktische Implikationen für Crowdlending Plattformen und Investoren diskutiert.

2 Theoretischer Hintergrund: Moral Disengagement

Zur Analyse des Einflusses von Moral auf die Investitionsentscheidung auf Crowdlending-Plattformen wurden zunächst die vier Dimensionen des Moral Disengagement Frameworks nach Bandura [6, 7] analysiert.

2.1 Kognitive Rekonstruktion der Handlung

Durch moralische Rechtfertigung, wohlwollende Sprache und vorteilhafte Vergleiche lassen sich soziale Handlungen kognitiv rekonstruieren. Werden Handlungen moralisch gerechtfertigt, dann werden sie damit erklärt, dass sie einem grösseren Zweck in der Gesellschaft dienen. Dies erlaubt es, ursprünglich als unmoralisch empfundene Handlungen moralisch akzeptabel zu gestalten, da Individuen normalerweise nicht dazu neigen, ohne Rechtfertigung unmoralisch zu handeln [6, 7]. Oftmals werden in diesem Zusammenhang unmoralische Handlungen zur Wahrung der Ehre und des Rufes ausgeführt [6, 10]. Die verwendete Sprache formt die Denkmuster als Basis der Handlung [7]. Dabei zeigt sich, dass Menschen auf grausamere Art und Weise handeln, wenn die Handlung selbst sprachlich von negativen Assoziationen befreit wird [11, 12]. Als weiteres Instrument der beschönigenden Sprache wird der passive Sprachstil verwendet, wodurch die Verantwortung für die Tat vom Individuum abgelenkt wird [13]. Die Methode der Marginalisierung durch wohlwollende Sprache und vorteilhafte Vergleiche nutzt bewusst das Prinzip der Kontrasterhöhung. Hierbei wird das eigene Verhalten einer moralisch besonders verwerflichen Handlung gegenübergestellt, wodurch es weniger gravierend erscheint [14].

2.2 Verleugnung und Minimierung der Verantwortung

Die Disengagement Mechanik wirkt als Modifikator auf die rationale Verbindung zwischen verwerflicher Handlung und schädlicher Auswirkung im moralischen Selbstregulationsprozesses [7]. Ein wichtiger Treiber von unmoralischen Handlungen ist deren Akzeptanz durch eine vom Individuum anerkannte Autorität, wodurch die Verantwortung verschoben wird. Dies erlaubt die Übertragung der persönlichen Verantwortung, wodurch der Selbstverurteilungsmechanismus umgangen wird. Dieser Effekt wird weiter verstärkt, wenn neben den befehlenden Personen auch die eigenen Bezugspersonen dem Befehl folgen [15, 16]. Durch Arbeitsteilung kann die Verantwortung verteilt werden, da jeder Arbeitsschritt die Verantwortung für das Endprodukt minimiert. Die Aufmerksamkeit des Individuums wird somit auf die eigene operationale Effizient gelenkt, wodurch die moralische Komponente der eigenen Arbeit in den Hintergrund rückt. Bei der Übertragung der Entscheidungsgewalt auf eine Gruppe fühlt sich Einzelne nicht mehr verantwortlich für das Gesamtergebnis [17].

2.3 Verzerrung der Konsequenzen

Die antizipierten Konsequenzen einer Tat sind ein starker Treiber der Selbstkontrolle des Individuums. Aus diesem Grund kann durch die Nichtbeachtung oder Minimierung der Konsequenzen für das Opfer das eigene moralische Selbstregulationssystem umgangen werden. Dieses tritt nicht in Kraft, wenn die schädlichen Folgen der eigenen Handlung verzerrt, minimiert beziehungsweise ignoriert oder nicht geglaubt werden. Wird ein Individuum mit den tatsächlichen Folgen konfrontiert, so diskreditiert es oftmals die Glaubwürdigkeit der Quelle. Aus diesem Grund ist dieser Mechanismus effektiver, je indirekter und damit unpersönlicher und weniger deutlich die Folgen wahrgenommen werden [6, 7].

2.4 Verantwortungsübertragung und Dehumanisierung des Opfers

Individuen neigen zur Selbstverurteilung, wenn sie schädliche Handlungen gegenüber Opfern durchführen, denen sie sich verbunden fühlen [6, 18]. Man unterscheidet dabei zwischen Schuldzuweisung und Entmenschlichung. Wird das Opfer oder die Umstände für die Handlung verantwortlich gemacht, kann sich der Handelnde auf das Recht zur Selbstverteidigung berufen. Durch Verdrehung der Kausalkette wird das Opfer für sein eigenes Leid selbst verantwortlich gemacht, was die Handlung entschuldbar macht und ein Gefühl der Selbstgerechtigkeit hervorruft [6, 7]. Ein wichtiger Treiber moralischen Verhaltens ist Gemeinschaftssinn. Die Vermenschlichung der Opfer führt zu einer deutlich reduzierten Aggressivität der Täter. Gleichzeitig hat die Entmenschlichung des Opfers eine entlastende Funktion für den Täter und reduziert die Wirkung des Selbstregulationsmechanismus [6, 7, 19].

3 Moral Disengagement in Crowdlending

Crowdfunding subsummiert verschiedene Formen der Finanzierung, die sich in ihrem Wertversprechen unterscheiden [20]: Belohnungs-basiertes Crowdfunding, Spenden-basiertes Crowdfunding, Eigenkapital-basiertes Crowdfunding sowie Kredit-basiertes Crowdfunding. In diesem Forschungsbeitrag wird das Kredit-basierte Crowdfunding untersucht, da es mit 11,08 Milliarden US-Dollar Kreditvolumen in 2014 eindeutig die finanziell relevanteste Form des Crowdfundings darstellt. Eigenkapitalbasierte Crowdfunding (USD 1,1 Mrd.) und hybrides Crowdfunding, eine Mischung verschiedener Modelle (USD 587 Mio.) folgen nur mit grossem Abstand [1].

Das gemeinsame Merkmal von Crowdfunding ist der öffentliche Aufruf im Internet an die „Crowd", die Summe der Nutzer der Crowdfunding Plattform. Eine Vielzahl an Investoren bringt dabei jeweils eine zumeist kleinere Summe ein, wodurch sich die Finanzierung auf eine grössere Anzahl an Investoren mit individuell kleineren Beiträgen verteilt.

Es konnte gezeigt werden, dass ein Zusammenhang zwischen Motivation und Wahrnehmung besteht, wobei Motive die Wahrnehmung gezielt lenken können um die gewünschte Handlung herbeizuführen. Dies wird durch eine gezielte Ausblendung von Informationen erreicht, welche der eigenen Motivation entgegenstehen. Gleichzeitig wird gezielt die Wahrnehmung auf Informationen gelenkt, welche die Motivation unterstützen [21].

Bezogen auf den Zusammenhang zwischen Motiven für finanziellen Gewinn und Moral Disengagement lässt sich dadurch ableiten, dass diese Steuerung der Wahrnehmung Einfluss auf das Moral Disengagement ausübt. Folglich bedeutet eine hohe Motivation zu finanziellem Gewinn, dass selektiv Informationen ausgeblendet werden, welche einer Optimierung der Rendite im Weg stehen. Dies können unter anderem Informationen zu moralisch bedeutsamen Themen sein, wie beispielsweise die negativen Auswirkungen für den Kreditnehmer [21]. Ebenso kann die gezielte Fokussierung auf finanziell relevante Informationen dazu führen, dass Investoren dazu neigen, das Verhalten ihrer Peers nachzuahmen. In diesem Fall wird die Verantwortung auf andere Investoren verteilt. Dabei wird die individuelle Verantwortung abgelehnt, da unterbewusst angenommen wird, dass Peers, die bereits in dem Projekt investiert sind, die Prüfung des Projekts zuvor schon vorgenommen haben. Durch diesen Mechanismus bleibt die eigene Unzulänglichkeit in der moralischen Prüfung unbewusst.

Hypothese 1a: *„Motive für finanziellen Gewinn begünstigen Moral Disengagement"*

Moralische Identität beschreibt das Eigenbild einer Person, welche ein relativ stabiles Selbstverständnis des Individuums ist, das sich aus spezifischen moralischen Charaktereigenschaften zusammensetzt. Diese formen hierarchisch geordnete Identitäten [22]. Eine Person, welche die eigene moralische Identität als sehr bedeutend wahrnimmt, fühlt stärkeres Mitgefühl für Mitmenschen auch ausserhalb der eigenen sozialen Gruppe. Aufgrund dieser Erkenntnis kann abgeleitet werden, dass Personen mit einer starken moralischen Identität weniger dazu neigen, unmoralische Investitionsentscheidungen zu treffen, indem sie Moral Disengagement Mechanismen aktivieren.

Hypothese 1b: *„Moralische Identität verringert den Einsatz von Moral Disengagement Mechanismen"*

Individuen treffen egoistische Entscheidung und können dabei gleichzeitig der Auffassung sein können, dass diese Handlungen korrekt sind [21]. Investoren stehen oftmals unter grossem Druck, eine Rendite zu erzielen. Aus diesem Grund ist anzunehmen, dass sich Investoren potentiell oftmals zwischen den Renditeansprüchen und den eigenen Moralvorstellungen entscheiden müssen. Dieses moralische Dilemma lässt sich kognitiv durch Moral Disengagement auflösen. Gleichzeitig rechtfertigt der Investor durch diesen Mechanismus unmoralische Entscheidungen, wodurch seine Prädisposition hierfür zunimmt.

Hypothese 2: *„Moral Disengagement der Crowdinvestoren ist positiv mit der Tendenz korreliert, unmoralische Investitionsentscheidungen zu treffen"*

Diese Arbeit argumentiert, dass verstärktes Moral Disengagement das vom Individuum empfundene Spannungsfeld zwischen eigener Erwartungshaltung an die Einhaltung moralischer Richtlinien und der tatsächlichen Handlung mindert.

Hypothese 3a: *„Moral Disengagement mediiert die positive Beziehung zwischen Motiven für finanziellen Gewinn und der Tendenz, unmoralische Investitionsentscheidungen zu treffen"*

Hypothese 3b: *„Moral Disengagement mediiert die positive Beziehung zwischen Moralischer Identität und der Tendenz, unmoralische Investitionsentscheidungen zu treffen"*

Folgende Darstellung illustriert das abgeleitete Mediationsmodell.

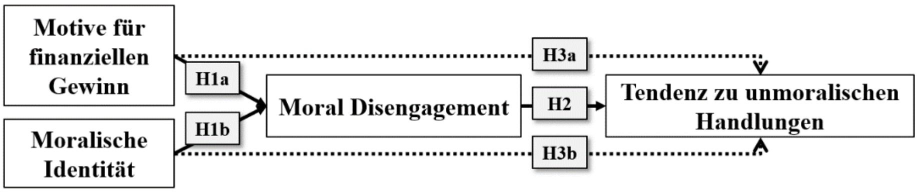


Abbildung 1: Darstellung des Mediationsmodells

4 Forschungsmethode

4.1 Datenerhebung

Um die Hypothesen zu testen wurde eine Online-Umfrage entwickelt, die sich an der Methodik von Baron et al. [21] orientiert und die Items bisheriger Moral Disengagement Studien adaptiert [21, 22]. Zur Messung der Tendenz zu unmoralischen Investitionsentscheidungen wurden ein Set aus vier Fallfragen (Vignetten) entwickelt, welche anhand einer kurzen Szenario-Beschreibung die Einstellung des Untersuchungsteilnehmers messen. Diese Szenarien im Rahmen der faktoriellen Umfragen erlauben es analog zu den Studien von Detert et al. [22] und Baron et al.[21], die Tendenz zu unmoralischen Handlungen zu messen. Dies erlaubt eine nuancierte Untersuchung der

Einstellung des Teilnehmers durch eine erweiterte Informationsdarstellung. Der Fragebogen wurde mit 10 Doktoranten vorgetestet, um sicherzustellen, dass die Fragen nicht zweideutig aufgefasst werden. Die gewählten Vignetten können online abgerufen werden[1], um das Untersuchungsdesign vollständig nachvollziehen zu können.

Die Umfrage richtete sich an Menschen mit Kenntnis oder persönlicher Erfahrung zum Thema Crowdfunding und wurde deshalb an Universitäten und in Crowdlending-fokussierten Foren verteilt. Total wurden 187 Antwortbögen des Onlineformulars eingereicht. Insgesamt 69 dieser Fragebögen wurden nur unvollständig beantwortet (mehr als 10% der Fragen wurden nicht beantwortet), wodurch sich die Zahl der Teilnehmer auf 118 reduziert. Die fehlenden Werte wurden mithilfe des Expectation-Maximization-Algorithmus (EM) vervollständigt, dessen Eignung mit dem Little MCAR Test festgestellt wurde. Die Umfrage wurde im Zeitraum Februar bis April 2016 durchgeführt. Die Teilnehmer sind mehrheitlich männlich (80; 67,8%). Der überwiegende Teil der Teilnehmer ist zwischen 18 und 24 Jahren (55,1%) beziehungsweise 25 und 29 Jahren (36,4%) alt. Dies erklärt sich aus der direkten Ansprache von Studenten, die meist in diese Altersgruppen fallen. 96% der Teilnehmer verfügen mindestens über die Hochschulreife, wobei 46,6% einen Bachelor und 11,9 einen Masterabschluss erreicht haben.

Die Aussagen wurden mithilfe einer 5-stufigen Likert-Skala abgefragt, wobei die Antwortmöglichkeit 1 „Stimme ganz und gar nicht zu“ und die Antwortmöglichkeit 5 „Stimme voll und ganz zu“ entsprechen.

4.2 Mediationsanalyse

Zur Untersuchung des Einflusses Moral Disengagement Mechanismen hat sich die Mediationsanalyse in der Forschung als robust herausgestellt [21, 22]. Die Mediationshypothesen werden anhand eines SPSS Makros [23, 24] geprüft.

Es lassen sich insgesamt drei Typen der Mediation unterscheiden: Komplementäre, Kompetitive und Nur-Indirekte Mediation [25]. Bei komplementärer Mediation und kompetitiver Mediation ist der direkte Effekt gegeben, allerdings deutet dies auf ein unvollständiges theoretisches Framework hin, da die Möglichkeit besteht, dass ein ausgelassener Mediator im direkten Pfad existiert. Ist der direkte Effekt allerdings nicht signifikant, so wird von nur-indirekter Mediation gesprochen. Hierbei ist der identifizierte Mediator konsistent mit dem theoretischen Framework und die Wahrscheinlichkeit, dass ein weiterer Mediator existiert, gering.

Preacher und Hayes [23] empfehlen den Bootstrapping Test des indirekten Effekts, um die Signifikanz des indirekten Pfades zu messen, da dieser im Gegensatz zum traditionell angewandten Sobel z-Test die Schiefverteilung der Stichprobe berücksichtigt [26]. Diese Arbeit folgt dieser Argumentation.

[1] Link: http://bit.ly/DS_MorDis_Fragenkatalog

5 Ergebnisse

Die Analyse mittels eines Mediations-spezifischen Makros [23, 24] wurde in SPSS durchgeführt und kann mit dem SAV-Datensatz[2] nachvollzogen werden. Die Variablen werden zuerst mit Hilfe des Cronbach Alpha Tests und der korrigierten Item-Skala-Korrelation bereinigt. Dabei zeigt sich, dass alle Variablen Alpha deutlich über den empfohlenen Grenzwerten von 0,6 und 0,7 für das Cronbach Alpha liegen [23, 24] und damit auch über vergleichbaren Werten von faktoriellen Umfragen im Bereich der Ethik und Moral [21, 22], was die Robustheit des Fragebogens bestätigt.

Die Analyse der zwei Mediationsmodelle bestätigen H1a, H2, H3a und H3b (siehe Abbildung 2).

H1a (angenommen): Die Analyse beweist, dass die Motivation zu finanziellem Gewinn die Neigung zu Moral Disengagement (Mediator) signifikant beeinflusst. Ebenso wirkt der Mediator signifikant auf die Tendenz, unmoralische Handlungen durchzuführen. Allerdings kann kein direkter Effekt nachgewiesen werden. Aus diesem Grund kann angenommen werden, dass eine nur-indirekte Mediation vorliegt. Der Bootstrapping Test ergibt ein BootLLCI von 0,009 (untere Grenze) und ein BootULCI von 0,124 (obere Grenze). Da dies nicht die Zahl 0 einschliesst, kann die indirekte Mediation bestätigt werden.

H1b (angenommen): Auch hier kann nachgewiesen werden, dass die Antezedens Variable (hier: Moralische Identität) einen signifikanten Einfluss auf den Mediator hat. Ebenso kann gezeigt werden, dass der Mediator signifikant auf die Neigung zu unmoralische Handlungen wirkt. Gleichzeitig liegt kein signifikanter direkter Effekt vor.

Der Bootstrapping Test ergibt ein BootLLCI von 0,024 (untere Grenze) und ein BootULCI von 0,213 (obere Grenze) und schliesst damit nicht die Zahl 0 ein, womit eine indirekte Mediation vorliegt.

Die Erkenntnis, dass eine starke moralische Identität ebenfalls Moral Disengagement begünstigt widerspricht bisherigen Erkenntnissen [22, 27] und zeigt auf, dass Individuen mit höheren moralischen Standards eher auf Moral Disengagement Mechanismen angewiesen sind, um unmoralische Handlungen zu begehen.

H2 (angenommen): Es konnte aufgezeigt werden, dass Moral Disengagement positiv mit unmoralischen Investitionsentscheidungen korreliert ist. Investoren stehen meist unter grossem Druck, eine grösstmögliche Rendite erzielen zu können. Dies kann dazu führen, dass eigene Moralvorstellungen übergangen werden müssen. Die Ergebnisse deuten darauf hin, dass sich dieses moralische Dilemma kognitiv durch Moral Disengagement auflösen lässt.

H3a (angenommen): Eine zentrale Annahme ist, dass Moral Disengagement als Mediator wirkt zwischen Motive für finanziellen Gewinn und der Tendenz, unmoralische Investitionsentscheidungen zu treffen. Dies konnte in der Mediationsanalyse bestätigt werden. Es handelt sich hierbei um indirekte Mediation nach Zhao et al. [26].

[2] Link: http://bit.ly/Dataset_DSchwarz

Damit ist der identifizierte Mediator konsistent mit dem vermuteten theoretischen Framework, Moral Disengagement kann somit als Mediator bestätigt werden.

H3b (angenommen): Die zweite entscheidende Hypothese besagt, dass Moralische Identität und die Tendenz, unmoralische Investitionsentscheidungen zu treffen, durch Moral Disengagement mediiert wird. Die Mediationsanalyse zeigt, dass eine indirekte Mediation vorliegt, welche den vermuteten Mediator bestätigt. Damit kann Moral Disengagement als Mediator im verwendeten Framework bewiesen werden.

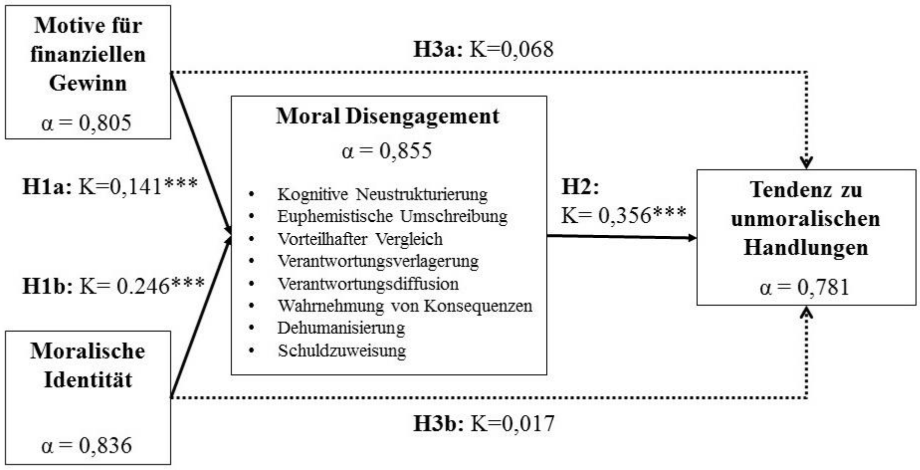


Abbildung 2: Ergebnisse des Mediationsmodells[3]

6 Diskussion & Implikationen

Crowdlending ist in den letzten Jahren eine ernstzunehmende Säule der privaten und geschäftlichen Kreditvergabe geworden. Durch die zunehmende Marktreife und die Rolle der Plattformen als vertrauenswürdige Intermediäre werden mittlerweile auch neue Investorengruppen erschlossen. Je relevanter für das Kreditgeschäft an sich und je gesellschaftlich akzeptierter Crowdlending wird, desto wichtiger wird es von akademischer Seite, die psychologischen Mechanismen zu verstehen, welche die Investitionsentscheidungen beeinflussen. Es konnte eine Forschungslücke im Bereich des Moral Disengagements identifiziert werden, die in dieser Arbeit anhand einer Mediationsanalyse untersucht wurde.

Die Arbeit präsentiert zwei massgebende theoretische Beiträge. Erstens, das in der Arbeit präsentierte Mediationsmodell fördert die Erkenntnisgewinnung zu moralischem Verhalten und Entscheidungsfindung im Crowdfunding. Dabei wurde der Fokus auf Moral Disengagement als holistischer Mediator gelegt. Diese Betrachtungsweise folgt dabei den Erkenntnissen von [6, 22]. Moral Disengagement und die Frage nach

[3] **Legende:** K = Koeffizient; α = Cronbach's Alpha; *** = p<0,01 | ** = p<0,05 | * = p<0.1

moralischem Verhalten wurde bislang in der Forschung zum Thema Crowdfunding weitestgehend vernachlässigt. Der Druck zu schnellen Investitionsentscheidungen mit limitierten Informationen und der Ausblick auf attraktive Renditen fördert beim Kapitalgeber die Tendenz zu unmoralischem Investitionsverhalten und somit das Umgehen des Selbstregulierungsmechanismus. Zweitens, stellt die Anwendung von Moral Disengagement auf die Crowdfunding Domäne eine Erweiterung der Forschung zum Moral Disengagement dar. Während Moral Disengagement in zahlreichen anderen Kontexten untersucht wurde, ist die Übertragung auf den Kontext Crowdlending neu. Dies dient als Ausgangslage für die weitere Untersuchung des Phänomens auch in anderen Crowdfunding-Arten dar.

Ausserdem präsentiert die Arbeit interessante Erkenntnisse für Praktiker. Das Verständnis über die Profitorientierung von Kapitalgebern und ihrer Tendenz zum Umgehen der moralischen Selbstregulierung ist sowohl für existierende als auch künftige Betreiber von Crowdlending-Plattformen von grossen Nutzen. Die Ergebnisse der Arbeit ermöglichen es Ihnen, Strategien und Mechanismen zur Vermeidung von unmoralischem Verhaltens von Kapitalgebern zu entwickeln. Diese sind notwendig um künftige Reputationsschäden zu vermeiden und damit zu verhindern, dass Kapitalgeber und Kapitalnehmer Crowdlending den Rücken zukehren oder womöglich eine härtere Regulierung des Marktes für Crowdlending ausgelöst wird.

Im Gegensatz zu Detert et al. [22] konnte nicht nachgewiesen werden, dass Individuen mit einer ausgeprägten moralischen Identität grundsätzlich weniger anfällig für Moral Disengagement Mechanismen sind. In dieser Studie wurde eine positive Korrelation zwischen Moralischer Identität und Moral Disengagement nachgewiesen. Dies könnte darauf schliessen, dass Individuen, denen moralisches Handeln besonders wichtig ist, zwangsläufig stärker auf Moral Disengagement Mechanismen zurückgreifen müssen um unmoralische Handlungen überhaupt erst zu ermöglichen. Umgekehrt kann argumentiert werden, dass Personen mit geringeren Ansprüchen an sich selbst hinsichtlich moralischer Standards auch nicht auf Moral Disengagement Mechanismen angewiesen sind, da sie eventuell die Handlung aus ihrer subjektiven Sicht nicht als unmoralisch empfinden. Diese Erkenntnis schafft die Grundlage für weitere Forschung hinsichtlich des Zusammenhangs von Moralischer Identität und Moral Disengagements.

Die Ergebnisse der Studie legen nahe, dass aufgrund der signifikanten Tendenz zu Moral Disengagement die Gestaltung des Kreditgesuchs einen entscheidenden Einfluss auf die Investitionsentscheidungen der Investoren hat. Für diese haben die Erkenntnisse dieser Studie insofern Einfluss, als dass sie Bewusstsein für das Thema und die Problematik von Moral Disengagement Mechanismen schafft. Hier gilt es zu überprüfen, welche Möglichkeiten bestehen, den Fokus der Investoren von Gewinnorientierung hin zu moralisch vertretbarem Investitionsverhalten zu lenken. So könnte untersucht werden, ob die Anzeige der anderen Investoren des Kreditgesuchs (oftmals mitsamt Profilbild) zu einer Minimierung der Verantwortung führt, da der Investor argumentieren kann, dass bereits andere das Angebot moralisch geprüft und als akzeptabel empfunden haben. Ebenso liegt eine Untersuchung darüber nahe, welche spezifischen Informationen über den Kreditsuchenden (z.B. Alter, Geschlecht, Kredithistorie, Verwendungszweck des Kredits, etc.) das Moral Disengagement beeinflussen. Auch von Relevanz ist die Darstellung der Kreditkonditionen. Hier kann vermutet werden, dass beispielsweise eine vergleichende Darstellung mit marktüblichen Konditionen das Risiko senken kann, dass Investoren Unwissenheit über die unvorteilhaften Konditionen eines spezifischen Kredits für Moral Disengagement einsetzen. Aufgrund der Neuartigkeit des Themas existiert noch keine weitere Forschung zu Moral Disengagement bei Crowdlending.

Ein wichtiger Beitrag dieser Forschungsarbeit ist folglich, implizit wirkende Mechanismen explizit zu machen. Es kann vermutet werden, dass das Wissen um die eigene Anfälligkeit für Moral Disengagement die Investitionsentscheidungen beeinflussen kann. Moralische Kontrollmechanismen müssen deshalb eine zentrale Rolle in der Entwicklung einer Crowdlending Plattform spielen.

Literaturverzeichnis

1. Massolution: 2015CF Crowdfunding Industry Report. (2015)
2. Bretschneider, U., Knaub, K., Wieck, E.: Motivations for Crowdfunding: What Drives the Crowd to Invest in Start-ups? In: 22nd European Conference on Information Systems (ECIS 2014). (Year)
3. Klafft, M.: Online peer-to-peer lending: a lenders' perspective. In: Proceedings of the International Conference on E-Learning, E-Business, Enterprise Information Systems, and E-Government, EEE, pp. 371-375. (Year)
4. Morse, A.: Peer-to-Peer Crowdfunding: Information and the Potential for Disruption in Consumer Lending. National Bureau of Economic Research (2015)
5. Schlechtriem, P.: Rechtsprechungsübersicht zum Bereicherungsrecht (Teil 1). Juristen Zeitung 39, 509-516 (1984)
6. Bandura, A., Barbaranelli, C., Caprara, G.V., Pastorelli, C.: Mechanisms of moral disengagement in the exercise of moral agency. Journal of personality and social psychology 71, 364 (1996)
7. Bandura, A.: Moral disengagement in the perpetration of inhumanities. Personality and social psychology review 3, 193-209 (1999)
8. Freedman, S., Jin, G.Z.: Do social networks solve information problems for peer-to-peer lending? evidence from prosper. com. (2008)
9. Collier, B.C., Hampshire, R.: Sending mixed signals: Multilevel reputation effects in peer-to-peer lending markets. In: Proceedings of the 2010 ACM conference on Computer supported cooperative work, pp. 197-206. (Year)
10. Cohen, D., Nisbett, R.E.: Self-protection and the culture of honor: Explaining southern violence. Personality and Social Psychology Bulletin 20, 551-567 (1994)
11. Diener, E., Dineen, J., Endresen, K., Beaman, A.L., Fraser, S.C.: Effects of altered responsibility, cognitive set, and modeling on physical aggression and deindividuation. Journal of Personality and Social Psychology 31, 328 (1975)
12. Gambino, R.: Watergate lingo: A language of non-responsibility. Freedom at Issue 22, 15-17 (1973)
13. Bolinger, D.: The loaded weapon: The use and abuse of language today. L.: Longman (1982)
14. Nisbett, R.E., Ross, L.: Human Inference: Strategies and Shortcomings of Social Judgment. {P}rentice-{H}all (1980)

15. Diener, E.: Deindividuation: Causes and consequences. Social Behavior and Personality: an international journal 5, 143-155 (1977)
16. Milgram, S.: Obedience to Authority: An Experimental View. Tavistock (1974)
17. Bandura, A., Underwood, B., Fromson, M.E.: Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. Journal of Research in Personality 9, 253-269 (1975)
18. Bandura, A.: Social cognitive theory of social referencing. Social referencing and the social construction of reality in infancy, pp. 175-208. Springer (1992)
19. Bandura, A., Caprara, G.V., Barbaranelli, C., Pastorelli, C., Regalia, C.: Sociocognitive self-regulatory mechanisms governing transgressive behavior. Journal of personality and social psychology 80, 125 (2001)
20. Haas, P., Blohm, I., Leimeister, J.M.: An Empirical Taxonomy of Crowdfunding Intermediaries. In: International Conference on Information Systems (ICIS). (Year)
21. Baron, R.A., Zhao, H., Miao, Q.: Personal Motives, Moral Disengagement, and Unethical Decisions by Entrepreneurs: Cognitive Mechanisms on the "Slippery Slope". Journal of Business Ethics 128, 107-118 (2015)
22. Detert, J.R., Treviño, L.K., Sweitzer, V.L.: Moral disengagement in ethical decision making: a study of antecedents and outcomes. Journal of Applied Psychology 93, 374 (2008)
23. Preacher, K.J., Hayes, A.F.: Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior research methods 40, 879-891 (2008)
24. Hayes, A.F.: Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. Guilford Press (2013)
25. Baron, R.M., Kenny, D.A.: The moderator--mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of personality and social psychology 51, 1173 (1986)
26. Zhao, X., Lynch, J.G., Chen, Q.: Reconsidering Baron and Kenny: Myths and truths about mediation analysis. Journal of consumer research 37, 197-206 (2010)
27. Aquino, K., Reed, A., Thau, S., Freeman, D.: A grotesque and dark beauty: How moral identity and mechanisms of moral disengagement influence cognitive and emotional reactions to war. Journal of Experimental Social Psychology 43, 385-392 (2007)

Examining the role of changing organizational culture on IT employees' commitment in M&A – Insights of a case study with a German software company

Aybala Ilgili

SAP Deutschland AG & Co. KG, Customer Engagement, Walldorf, Germany
ay.ilgili@sap.com

Andreas Eckhardt

German Graduate School of Management & Law, Heilbronn, Germany
andreas.eckhardt@ggs.de

Abstract. Managing IT employees' commitment in and after mergers and acquisitions (M&A) in the software industry is a key reason for success and failure of these initiatives. Thus, the purpose of this research is to identify influencing factors for IT employees' organizational commitment level during this process. Based on the case of a German software company that has been involved in several M&A over the last couple of years, qualitative data was gathered via personal semi-structured interviews with acquired employees, managers, and employees concerned with the mergers. All interviews are conducted and evaluated as well as presented using a narrative case approach. Results of the interviews suggest an overview on fields of action and processes in the field of organizational culture. Findings show that those IT employees, who're feeling being part of the new organization are more likely to stay motivated and committed while creating something new. In contrast, a strong culture at the prior employer lowers the feeling of being part of the new software company. As main contribution, this research discloses the two-sided coin of communication leading both to enhanced commitment due to increased transparency as well as to information overload decreasing employees' commitment during an M&A.

Keywords: Employee Commitment, Software Industry, Mergers and Acquisitions, Case Study

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Ilgili, A.; Eckhardt, A. (2017): Examining the role of changing organizational culture on IT employees' commitment in M&A – Insights of a case study with a German software company, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1513-1525

1 Introduction

Researchers believe that companies engage with mergers and acquisitions (M&A) in order to increase profitability. But many other reasons, such as operating and financial synergies, contribute to an increasing activity of M&A as well [1]. Especially in the software industry, the progressive level of competition linked to M&A is remarkable. M&A trail organizational changes, which lead to different feelings, such as frustration, insecurity, stress, or non-appreciation among employees. These reactions can have a huge impact and cause the employees to rethink their commitment level to the organization. Although M&A might have many gains, such as increased turnover or profit, some companies look good on paper in the beginning, but hick-ups will confront both sides. M&A, especially in the IT industry, are very trendy and much sought-after since 2005, because these unions increase sales performance, grow assets, and develop networks.

While M&A are in great demand, companies have different processes, approaches or methods regarding the management of their employees during M&A, and this is an often critically discussed and considerably researched topic [1-3]. Within the Information Systems (IS) field, studies on IT workers' commitment have a long tradition, but to date there is no work examining their commitment during an M&A initiative [4-7]. Prior research only emphasizes that IS managers need to focus on the effect of motivating potential, group cohesiveness, and stress level during change initiatives [4]. According to Klaus et al., these aspects contribute to greater IT worker commitment, minimize turnover, increase their retention, which, especially in the IS industry, contributes to the company's profitability and long-term survival during larger organizational changes [4-5]. While employee commitment is considered being very important and the core variable in several research approaches is IT workers' turnover [6-7], there is still a research gap when it comes to the role of employee commitment during different stages of organizational development, such as after a merger or an organizational acquisition.

Therefore, this research aims to answer the following research questions:

RQ1: What are drivers and consequences for the commitment of acquired IT workers during an M&A initiative?

RQ2: What is the role of the changing organizational culture before, during, and after the M&A initiative?

In order to answer these research questions, this paper applies a qualitative research approach and introduces cases of several M&A processes involving a major German software corporation and acquired firms. In order to gather useful information, semi-structured personal interviews with managers, and employees affected by M&A were conducted. The respective data was evaluated and presented, using a narrative case approach [8]. Before we represent our case study results, we review prior literature for factors and concepts influencing employee commitment during M&A initiatives in the following section.

2 Research Background

This section will share theories and concepts as well as categorize topics on the role of organizational culture possibly affecting employee commitment during M&A. These aspects show the cause behind the proposition formed at the end of this section, which state the relevance of this study.

2.1 Influence of Old and New Organizational Culture during M&A

Cartwright and Cooper compare M&A with a civil marriage of two people and explain that they disturb the cultural peace, because when two different cultures come together, it might clash [9]. The bigger the difference, the bigger is the clash. Robinson indicates that start-up companies shoot from the hip while large concerns have their established processes [10]. Both have their advantages, but this can cause an organizational culture collision very often and lead to blurry working environments, conflicts among the employees and stress. Thus, less employee commitment is given and consequently the organizational performance is affected by job dissatisfaction [9].

Although most organizations agree on the fact that culture creates value to an organization, they have not found a successful recipe, because there is no 'right' or 'perfect' culture. It only depends on how two cultures get along with each other, just like in a marriage [9].

Cultural Orientation and Types

Previous literature on organizational behavior provides the foundation for Cameron's "Competing Values Framework" consisting of four quadrants – collaborate, control, create and compete [11]. Each one stands for a particular type of culture. The create-oriented culture has its focus on the products and future innovations, constantly developing new ideas based on technology and processes. Employee development, involvement and satisfaction are the characteristics of a collaborate-oriented culture with the aim to build skills and employee commitment. The control-oriented culture concentrates on transferring better processes to improve the efficiency. Whereas the compete-oriented culture strictly follows and observes the external competitiveness, which is determined by shareholder value, customer satisfaction, market shares, and so on [11]. An organization might have dominating aspects in its culture, but usually a little bit of every quadrant is present, because of differing forms of value creation. This might influence the merger type, because in an M&A the cultures can differ dramatically due to two different cultural approaches. This can result in frustrated employees leaving the organization or not performing efficiently in all merger phases.

Communication Culture

Regardless which cultural orientation or type companies have, employees are involved and will behave accordingly to the alignment of two organizational cultures. Larsen explains that in order to have a good relationship with employees and establish trust in general, a transparent communication is needed [12]. Furthermore, Larsen suggests supporting the communication strategy with managerial listening sessions [12]. Listening people evokes a feeling of appreciation and values opinion among the employees. This triggers loyal behavior towards the organization, lowers fear and increases satisfaction. Watson emphasizes that fear, doubt, and uncertainties are steady companions of M&A, because people will always have the question of what will happen next [13]. Furthermore, Watson stresses that an honest, informing, engaging and balanced communication will not result in a high level of fear [13]. Adding to this, Nahass et al. point out that communication has to be present at all stages – from the beginning to the end and long after the integration – because this transparency takes away fear, builds trust and therefore increases employees' commitment [14]. On this basis, Schwaab et al. illustrate that adequate information for employees' motivation is necessary, but not enough [14]. The top management has to show also persuasiveness, assertiveness, credibility and even more. Therefore, people who are part of the M&A should receive and give enough information at each stage at the integration via communication. Every concern needs attention, in order to build trust, increase the bond strength, lower uncertainty and grow employee commitment.

Cultural Working Environment

Organizations have different working environments, which are an important part of the culture. During an M&A, employees are in an identifying mode if they are compatible with the new organizations' work environment. For example, many companies have dress codes, core times or modern and colorful offices as well as room for creativity and new ideas. Mitchell et al. stress that employees seeing themselves suited to the organizational culture, are more attached to the organization in a professional and personal way [16]. Hom and Griffeth and Kane-Urrabazo suggest that the strongest indicator for employee retention is the creation of a positive working place and the linkage to job satisfaction [17-18].

Cultural Awareness

Schwaab et al. explain that M&A cause stress and fear for the individuals, because they are personally affected [15]. Whether it is the loss of workplace, the new technology, a change in the team, the long sought for career step, or the new organizational culture with a possible new name of the company. Schwaab et al. stress that M&A will always trigger fear, because each individual has different experiences, personalities and emotions [15]. It is more important to figure out how to deal with this fear and stress level employees have. Hereby, the organizational environment and the culture have a major role. The culture can lower the fear or stress level of employees by seeing it as a chance for the new or creating diversity [15]. In order to understand this, organizations have to achieve this cultural awareness and spread it.

Organizational Development & Diversity

The organizational development is a major factor in a human resources structure and adds to the company's' culture. According to Groysberg and Fernandez-Araoz, an organization needs to follow three steps, in order to find and develop their rising stars.

First, create a clear structure for the organizational priorities. Second, know how to select the right high-potential employees and third, have a plan on how to develop and reward these high-potentials. These steps can also be applied during an integration, in order to keep the acquired employees on board and increase their commitment level [19].

2.2 Research Proposition

After discussing the issues of M&A and examining a part of the existing literature regarding managers, work task, organizational culture and individual differences in general and in M&A, the focus is now shifting to the aim of this case study. It is based on a qualitative data gathered in semi-structured interviews. As mentioned before, there is still a research gap when it comes to the role of employee commitment during different stages of organizational development. It appears that although many companies have a theoretical construct with an overall picture of influencing factors of present or future employee commitment, still M&A fail or the aftermath endures too long. Therefore, within this case we observe the following proposition.
Proposition: Cultural integration at the software company is not always successful, because communication and information can lead to confusion and consequently exclusion.

3 Methodology

The subsequent sections describe the methodology including the applied research approach, case description, and data collection instruments as well as data analytics instrumentation.

3.1 Research Approach

The methodology of this case study follows a concept named "*a linear but iterative process*". Yin outlines that a case study is *"a contemporary phenomenon (the "case") in its real-world context, especially when the boundaries between a phenomenon and context may not be clearly evident"*, whereof the researcher has little control over the phenomenon or its behavior as well as a contemporary focus [8]. This case study consists of Yin's five features. The interviews allowed the participants to share what they want and feel and were not limited to questionnaire procedures (people's performance under real life conditions). Additionally, the three different groups gather different participant views, which emphasize the feelings and values of people who actually lived the incident. The contextual conditions, such as the social and external environment, have a strong effect on these incidents that people lived. Furthermore, the company employees bring in new perspectives and exclude the branding experiences from the past (preexisting and originating concepts). In addition, the three different groups base its conclusions on various sources. These five features lead to a mosaic of methodological and philosophical views, which contribute to different interpretations of people of the same event [8]. The multiplicity in interpretations gives the opportunity to present how opinions differ, because every individual interprets and feels experiences – the same reality – in a different way [11]. Furthermore, the different perspectives show where the company has faulty processes and what leaves room for improvement in order to reach a high level of employees' commitment during and after M&A.

3.2 Case Description

The company in our case has about 78.000 employees in 130 different countries and acquires companies since 2007. To observe the effect of old and new organizational culture, the interviewers selected a total number of 20 individuals from the software company in Germany, which represent the larger group of employees regarding the M&A topic. These interviews demonstrate three expert groups, whereof eight interviewees were acquired employees from different acquisitions during the past thirteen years (AcqEmpl= 8), seven current employees concerned with M&A and HR processes (Empl= 7), and five employees holding management positions (Mgmt= 5). The acquired employees shared their views in regards to contextual environment such as communication, technology, security, manager's role, trust, pace, information flow, being a part of the old and new organization, as well as their commitment level during the integration. The current employees contributed by explaining the effect on their internal daily working tasks in regards to feedback of acquired employees. The management group accompanied integrations from different departments, such as labor relations, M&A, board office, as well as HR and shared their so-to-say external perspectives on integrations. Each member of every group was part of M&A directly or indirectly of approximately seven different acquisitions in the last thirteen years. This contributed to multiple perspectives and sources of evidence (triangulation) in regards to integrations. The length of each interview was between 20–45 minutes and had the format of an open and virtual conversation. In order to gather the data and record the interviews in mp4 format, quiet rooms at the workplace were provided. The interviewer recorded the sessions and transcribed them subsequently. The interviews were conducted in a rough time span of one month, absolute anonymous and on a voluntary basis. In total, there are eleven female and nine male interviewees.

3.3 Data Collection Instrument

The different interview groups as well as acquisitions of this case study reason a multiple case design categorization. It explores a unique case as a unit with multiple acquisitions and data including documents, interviews and physical artifacts in form of company servers and former M&A information as sources of evidence [20].
This case study leans on a descriptive and explanatory approach that pumps as much information as possible [20]. Therefore, the interviewer used semi-structured narrative interviews. The interviewer asked the participants consistently for precise emotional states during the integration in order to overcome the past time frame of 13 years and filter the experienced feelings. Thus, the interviewer was able to support the

participants' deviation or asked further questions in case of insufficient answers. This structure of interview is also called in-depth inquiry and was supported by "how", "why" and/or "please describe" questions as well as matters with scalable answer possibilities following the request to describe their decision [8]. Besides the interviews, notes, tables, and narratives created the case study database.

3.4 Data Analytics Instruments

In a qualitative research, the focus is on the relevance of reviewed subjects concerning content representativeness. This means, the individual aspects play a major role. For this, Yin explains that generalization, validity and reliability have to be considered entirely [8]. In this connection, several steps were listed – paraphrasing, thematic order and comparison, conceptualization and generalization. First, paraphrasing implies that the content of the interviews are reflected and statements to a certain subject are highlighted, following the main themes from the interviews were assigned and compared (thematic order). In the conceptualization phase, commonalities and differences considering the theoretical body of knowledge of this study are expressed. The final phase of the analysis is the analytical generalization aspect, which compares the previous theoretical propositions to the results of this study. This study includes interviewees with acquired employees, managers as well as current employees of the company comprising M&A of the last thirteen years and is exploring an empiric subject by following Yin's iterative procedures [8].

4 Results

This section delivers categorized insight into the different perspectives and aspects of cultural change before and after the integration. For this, the resulting data is described by examining and categorizing the feedback of 20 semi-structured interviews and the above mentioned literature regarding influencing factors for employees' commitment. The following themes originated from this procedure.

Old Culture

All of the acquired employees mentioned how comfortable and fully committed they felt with their old employer. Additionally, the start-up or hands-on mentality was emphasized. Due to several factors, their commitment changed once the integration to the software company took place.

"It was a great place to work. We had something called "Hybris Spirit", which kind of described our way of work…"It's a life style, not just a work place"… It was a culture open for mistakes. We had quite a hands-on mentality, because we did everything for the first time. It was quite a family feeling." (Interviewee1, AcqEmpl)

One pointed out that the average age in their former team was younger, which contributed to a more dynamic working environment and behavior. Interestingly,

some mentioned their loyalty towards their old employer, because the relationship with the organization was not only on a professional, but also on a personal level.
"…the things they've done for me - both professionally and personally - have been pretty great; so I feel a great deal of loyalty there." (Interviewee4, AcqEmpl)

Also, almost all of the acquired employees pointed out that their colleagues and 'creating something new together' contributed to a great deal of loyalty and a feeling of 'being an important part' of the company (Interviewee5, AcqEmpl, p. XXI). Interestingly, one explained that a start-up company has only a few employees who might have to do more than they had to. This includes the willingness to dedicate more time and effort to the organization on a voluntarily basis.
"Culture - especially - was a big, big topic… when you are kind of small, you need people to play many roles, and people that are willing to not simply do their stuff, but take on things that need to be done." (Interviewee6, AcqEmpl)

These aspects all led to high levels of commitment with the prior employer and prevented the development of commitment with the new employer.

Organizational Culture
Many of the interviewees highlighted that the company has a lot of diversity in terms of people, culture, and languages and is a very open-minded organization with room for new ideas.
"I am not native German… culture means that everyone is welcome. They don't expect me to speak German… to understand their culture, at least in a professional environment. Everything is diverse and respectful." (Interviewee11, Empl)

Furthermore, the culture of networking in terms of coffee appointments was mentioned in a very positive way. The coffee corners mentality offers the face-to-face meetings, which give employees a personal note to their work. Nevertheless, employees note the lack of understanding relating to processes and procedures among the acquired employees during the organizational change. It was mentioned that there was no room for flexibility to approach processes differently. Everyone had to follow the established process steps, which were time consuming and partly confusing. *"Since the company is such a big company all processes take rather long. Everything is kind of linked to a formal procedure. And employees from acquired companies, which are usually much smaller than this software company, are not used to all this formality which at times can lead to frustration." (Interviewee14, Empl)*

The Giant
Many interviewees mentioned that the size of the organization and the volume of the change matters when two cultures come together. Instruments such as buddies, coffee corner sessions or organizational tools provide an additional insight into both organizations. However it was largely expressed by acquired employees that the software company was perceived as the 'Giant'. The primary reason for that was the imposition of working behavior and neglecting their culture.

"It was more of 'Wow. We bought you. We're a huge company. We know what we are doing. Here's what you're going to do.' And for some things they're probably right to do that. In other case… the things they were trying to tell us… simply wouldn't work for us. We understand that the software company bought us and therefore they have the right to make certain decisions. But at the same time we're trying to tell them what made us successful…" (Interviewee4, AcqEmpl)

Interestingly, one interviewee explained that employees' commitment level decreases automatically in an integration with two companies differing this much in their sizes. In a start-up organization, this combination was seen as a given fact. Due to the size of the company, it was rather perceived as impossible to establish this kind of combination and therefore a high level of commitment in the company.
"I'm not working for and with, at least, 75 thousand people - I will never meet them, I will never talk to them. I have no idea what they are doing. And it's likewise. So I rely on a lot of things that happen without having the tiniest chance of influencing anything. By way of not being ... the 1/75000th part of a company, my potential influence on anything, of course, has diminished as well. And in that sense, of course, it's harder to build up that relationship and that amount of trust and commitment to a company that you had before." (Interviewee6, AcqEmpl)

Respect
Showing mutual respect when two organizations come together was found to be a fundamental element of culture. The managerial interviewees are aware of the fact that people at the company tend to think their working behavior is more structured compared to the small-acquired company. The correlation of the 'behemoth' and the 'small, innovative and shooting from the hip' company causes disturbance, because the basic behavior among individuals have to be given, in order to create something new together. Employees believed that respect needs to be given first and it is in everybody's own hands. Therefore, a respectful treatment and showing interest in the counterpart is a crucial aspect.
"There is also a cultural aspect, which is respecting the culture that they come from and respecting the success that they had as a company… When they come in, there is this tendency to want to impose the same structures on them as we have... I think, sometimes some of that is necessary… And if you do need to impose certain rules, than we need to do it in a more sensitive way. And I think the other cultural aspect is just valuing what they do well. Making sure that this is being recognized and appreciated." (Interviewee19, Mgmt)

Having a Voice
Current employees assumed that the new acquired employees don't have a voice although many instruments are offered. In particular, the coffee corner session was described slightly underemphasized that it might not be the best instrument, because not every person is comfortable with speaking out concerns in front of others. *"Sometimes they get a voice… I don't have the impression that the exchange is going on very well… Yes, they have some options to express their selves in coffee corner*

sessions. If they dare to, but I think not everyone likes to speak up in a coffee corner session and to express his/her concern. This depends on your character...” (Interviewee12, Empl)

Interviewees found that being a part of this company have a great influence on the colleagues, the emotions and the organization. The cultural integration is at an acceptable level, but still has room for improvement. In order to make everyone a part of the organization and avoid ‘us vs. them’ attitude, listening and responding to the employees is one starting measure in cultural workshops, considering that finding the balance in cultural differences varies case by case. Therefore, approaches have to constantly be customized.
“That's another one of those things that you need to figure out in the cultural assessment upfront. Is this the kind of company that would like that and wants to feel as a part of the software company? Or do they want to be left alone a little bit more. Often, they do want to be left alone. But then the question is, we don't want to leave them alone too much, because than they won’t feel like a part of the family. So, you always have to come up with that balance.” (Interviewee19, Mgmt)

5 Discussion

We started our qualitative investigation with the proposition that cultural integration in M&A is in the software industry is not always successful, because communication and information can lead to confusion and exclusion. We outline the major implications of our research on drivers and consequences of IT workers’ organizational commitment in M&A initiatives in this section. Therewith we emphasize how IS research on managing the workforce can benefit by our approach [4-7].

5.1 Implications for Research

In integrations, culture is multifaceted. One important aspect is the cultural orientation. Following Cameron’s explanation about cultural orientation, the software company belongs to the create-oriented section, since it has its focus on the products and future innovations, constantly developing new ideas based on technology and processes [11]. Relating to M&A and according to the theory by Cartwright and Cooper [9], the company follows the approach of a redesign merger during an integration, which includes creating win-lose situations. This means, the dominant acquiring company forces its culture onto the acquired company. According to many statements from interviewees, the software giant’s cultural dominance during an integration applies completely. One reason is the difference in the size of the companies. Most of the acquired organizations are significantly smaller than the software company. Another point is the lack of flexibility from the company’s side when it comes to processes and procedures. On the one hand, these established and structured activities are needed. On the other hand, acquired employees see the

company as a behemoth forcing their cultural aspects onto others. The expectations are rather different and asking for the acquired company's 'recipe' for success is missed.
Secondly, communication is essential for employees going through a complex organizational change. The interview results are the basis for dividing the negatively perceived communication and information in three categories. First, interviewees pointed out that lack of communication leads to uncertainty among the employees and consequently results in dissatisfaction. Second, the wrong communication/information leads to distrust, because too optimistic communicated aspects cannot be fulfilled later on. Whereas too much communication/information is very overwhelming and end up e.g. in deleting mails.
"Yes, the software company sends out tons of emails…, but there is always a delete button if you are not interested." (Interviewee1, AcqEmpl)

All three categories result in confusion, frustration, lack of trust and dissatisfaction, which are influencing factors for commitment. Although the software company tries to include the employees during an integration via their numerous communication instruments, employees are totally confused or overwhelmed at some point, which leads to exclusion. These aspects negatively affect employees' commitment level. Based on the mentioned literature, this could have a negative effect on employees' productivity and therefore the company's success. Moreover, this could be a reason why some employees are still rebuilding their relationship with the new employer, although the integration was several years ago.

5.2 Implications for Organizational Practice

The company observed in this case studies is known as a very attractive and reliable employer that puts a huge focus on their organizational culture and employees. The happier the employees, the better the performance. Moreover, the company values employee satisfaction and therefore puts their focus on motivation with various programs or methods. One program is called the Fellowship program, which is designed to engage employees in strategic projects and to grow strengths in different business areas. It can range from 4 weeks to 6 months in length and includes a possible relocation. The main purpose of this program is that employees accelerate readiness for broader responsibilities and collaborate on strategic projects to support execution of the company's strategy. One interviewee mentioned that there was an employee, who took part at this fellowship program, mainly to support them in their integration phase. His responsibility was to onboard his future colleagues in company systems and tools including its forecasting, budgeting and Q&A sessions. This helped to shorten the learning curve. Moreover, they felt appreciated and addressed all the concerns or issues in that area. The fellow on the other hand, was able to dive into a different department in another country. Therefore, this program is an ideal option to expand in the pre-merger phase of M&A in the future. This could help to reduce the frustration level among the employees regarding the technological area and provide a confident and positively emotional stage during the integration among the employees.

Furthermore, it would suit the acquired employees, because they often mentioned in the interviews that they prefer the face-to-face contact in such organizational changes.
A second point, which was mentioned by some interviewees, was related to the cultural workshops. Here, it was a group of employees, who visited the acquired companies to approach possible issues, offer transparency, discuss strategy, set up cultural scans, ask for feedback etc. As far as it was explained, this only happened one time during an M&A, although it received only positive feedback to this procedure. This might be another option to tackle and implement the cultural workshops as inherent part in the M&A processes in the pre-merger phase for future integrations.
There is a lot of energy in the beginning of an M&A and the organizational change will carry many emotional stages. It will start with excitement and continue with doubt and discomfort. At this turning point, difficulties with various aspects will rise. The importance is to use the right measures at this explicit turning point to reduce the negativity and increase motivation and commitment. Moreover, it is important to understand that acquisitions are a good thing and help to secure a company's success. As one manager explained:
"…these acquired companies, they help us and they secure our revenue stream, our success and with that we have the workplaces we have right now. For every acquired workplace, we secure another year or something like that and to me it's clear that it's that way, but maybe that should be clearer to everyone, or should be made clearer to everyone here." (Interviewee20, Mgmt)

5.3 Limitations and Interesting Avenues for Further Research

Nonetheless, the study has limitations. The study was conducted with 20 employees and based on qualitative data. Meaning, although the theory proves that lack of trust, communication etc. leads to frustration and therefore to low commitment level, the values from the interviews are not analog to the statements received from the interviewees. While acquired employees are very frustrated in many aspects, their commitment level is only minimally lower, which is a compliment for the new employer. Surprisingly, current employees and managers estimated the commitment level of acquired employees extremely lower. The same findings were examined with the overall expectations in relation to commitment level. We assumed that the company must have other positively influencing factors for commitment, which convince the employees and avoid the extreme decrease of commitment, e.g. the IT industry, size of the company, numerous opportunities. However, in order to find out the reasons for the nearly non-changing commitment level among employees during an integration, large-scale interviews in the IT industry might be an idea. Hereby, it is suggested to put the focus on qualitative data again, since it offered a good platform with much information.
Secondly, interviewing drop-offs of former M&A can support creating a wider understanding of this matter. This could give further insights to commitment, turnover, and other predictors that could add value to the organization and research.

References

1. DePamphilis, D. M.: Mergers, acquisitions, and other restructuring activities: An integrated approach to process, tools, cases, and solutions. Amsterdam: Academic Press (2013)
2. Faulkner, D., Teerikangas, S., & Joseph, R. J.: The handbook of mergers and acquisitions. Oxford: Oxford University Press (2012)
3. Wollersheim, J.: Dynamic Capabilities im Kontext von Mergers & Acquisitions: Erfolg von Zusammenschlüssen von Organisationseinheiten mit unterschiedlichen Routinen (1st ed.). Wiesbaden: Gabler (2010)
4. Klaus, T., LeRouge, C.,, Blanton, E.J.: An examination of the relationships between select nature of work characteristics and organizational commitment of IT professionals. SIGMIS CPR Conference, April 10-12 (2003)
5. McKnight, H. D.: Relational roots of IT worker organizational commitment. SIGMIS CPR Conference, May 28-30 (2009)
6. Eckhardt, A., Laumer, S., Maier, C., & Weitzel, T.: The effect of personality on IT personnel's job-related attitudes: Establishing a dispositional model of turnover intention across IT job types. Journal of Information Technology 31, 48-66 (2016)
7. Maier, C., Laumer, S., Eckhardt, A., & Weitzel, T.: Who really quits? A longitudinal analysis of voluntary turnover among IT personnel. The DATA BASE for Advances in Information Systems 46 (4), 26-47 (2015)
8. Yin, R. K.: Case study research: Design and methods (5th ed.). Los Angeles, Calif. [u.a.]: Sage Publ. (2014)
9. Cartwright, S., Cooper, C. L.: The role of culture compatibility in successful organizational marriage. Academy of Management 7, 57-70 (1993)
10. Robinson, T.: Corporate unions put the squeeze on networking professionals. Merger Mania, 27-32 (2005)
11. Cameron, K. S.: Competing values leadership: Creating value in organizations. Cheltenham, UK, Northampton, MA: E. Elgar Pub (2006)
12. Larsen, K. H.: Communication – The key to successful mergers & acquisitions? Aarhus University, Department of Language and Business Communication (2009)
13. Watson, L.: Addressing culture, communication & customer integration issues to avoid M&A failure. Peak Performance International (2011)
14. Nahass, G., Smith, J., Curragh, M., & Colombo, L.: M&A communications. Communicating to engage and motivate people throughout the deal. PwC Deals (2014)
15. Schwaab, M.-O., Frey, D., Hesse, J.: Fusionen: Herausforderungen für das Personalmanagement (1st ed.). Heidelberg: Verlag Recht und Wirtschaft (2003)
16. Mitchell, T. R., Holtom, B. C., Lee, T. W., Sablynski, C. J., & Erez, M.: Why people stay: Using job embeddedness to predict voluntary turnover. The Academy of Management Journal 44 (6), 1102-1121 (2001)
17. Hom, P. W., Griffeth, R. W.: Employee turnover. Cincinnati: South/Western (1995)
18. Kane-Urrabazo, C.: Management's role in shaping organizational culture. Journal of Nursing Management 14, 188-194 (2006)
19. Groysberg, B., Fernandez-Araoz, C.: How to hang on to your high potentials. Harvard Business Review, October (2011)
20. Rowley, J.: Using case study in research. Management Research News 25 (2002)

The Core Capabilities of Green Business Process Management – A Literature Review

João Carlos Maciel

University of Liechtenstein, Institute of Information Systems, Vaduz, Liechtenstein
joao.maciel@uni.li

Abstract. Environmental sustainability has become a key concern of contemporary organizations, due to the impact of their operations on the environment. Organizations can be seen as sets of business processes that aim to generate business value, and business process management can play a significant role in designing and implementing more environmentally sustainable processes. Green BPM is the discipline that integrates sustainability-related thinking into business process management. In this paper, we present the results from a literature review on Green Business Process Management. Using the six core elements of BPM – Strategic Alignment, Governance, Methods, Information Technology, People, and Culture – as an analytical lens, we identify key capabilities required in Green BPM. Based on our analysis, we derive a Green BPM lifecycle to identify relevant methods, techniques, and approaches that can be applied to help organizations design improved business processes that exert a minimal impact on the natural environment.

Keywords: Green BPM, Environmental Sustainability, Core Elements of BPM, Green BPM Lifecycle, Literature Review.

1 Introduction

Environmental Sustainability is one of the various concerns of contemporary organizations, and sustainability is now often seen as an explicit performance dimension, in addition to traditional ones such as cost, time, and quality [1]. Sustainability has become increasingly important as organizations need to comply with rules and policies with regards to their operations and their consequences on the natural environment [2].

Organizations can be seen as sets of business processes operating to generate business value. The enterprise's business processes play an important role in contributing to the carbon footprint an organization emits [3]. Therefore, as BPM focuses on the understanding and improvement of an enterprise's business processes [4], it can be a significant enabler towards more environmentally sustainable organizations. Furthermore, Green BPM is the discipline that integrates sustainability-related thinking into business process management.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Maciel, J.C. (2017): The Core Capabilities of Green Business Process Management - A Literature Review, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1526-1537

In this paper, we present the results from a literature review on Green Business Process Management. Prior literature reviews on the topic can be found [5], [6], [7]. However, this study takes into account a different approach to analyze the current state of Green BPM research. Using the six core elements of BPM [8] – Strategic Alignment, Governance, Methods, Information Technology, People, and Culture – as an analytical lens, we identify key capabilities required in Green BPM. This study also differs from previous literature reviews as it depicts a Green BPM lifecycle, which encompasses diverse methods and techniques derived from the literature. Our essential research question is:

What is the current state of Green BPM research?

The remainder of this paper is structured as follows. Section 2 provides an overview on the concepts of Green BPM and green business process, as well as on the six core elements of BPM. Section 3 describes the methodology used for the literature review. Section 4 presents the main findings from the analysis of the literature. Section 5 presents the Green BPM lifecycle derived from literature. Potentials for future research are disclosed in Section 6. Finally, Section 7 concludes with a summary of the results.

2 Research Background

2.1 What is Green BPM?

We firstly present the concepts of sustainability and BPM, which are followed by a definition of Green Business Process Management. Sustainability is commonly defined as the "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" [9]. Business process management can be defined as the discipline that provides appropriate concepts, methods, and tools to model, implement, operate, and monitor business processes [10]. Different yet similar definitions of Green BPM can be found. They commonly build on the business process lifecycle from "classical" BPM, and add the ecological factor. Examples are the "dedicated consideration paid to the environmental consequences of these business processes" [11], the optimization of business processes considering the ecological dimension [12], and the aim to support environmental objectives [13].

Thus, one can understand Green BPM as the discipline that provides appropriate concepts, methods, and tools to support business process modeling, implementation, execution, monitoring, and continuous change with dedicated consideration of the environmental impact of these business processes.

2.2 What is a Green Business Process?

The execution of a business process has always a certain impact on the environment [12]. Business processes play thus an important role in contributing to the carbon footprint of an organization [3]. As a business process transforms inputs into outputs, the renewable and non-renewable inputs as well as the type and environmental quality of the outputs determine the sustainability of a business process [14]. In this line of thought, the extent of renewable inputs as well as the extent of emissions (outputs) that can be kept within the assimilative capacity of the environment should be maximized, while the extent of non-renewable inputs and the extent of emissions (outputs) that exceed the assimilative capacity of the environment should be minimized [14].

Therefore, a green business process is an environmentally conscious business process that is necessary, efficient, effective, agile, and measurable in the context of an organization [3] and delivers organizational value with a minimal impact on the natural environment [14]. Additionally, from a more technological perspective, an energy-aware business process is a service-based business process that considers functional and non-functional requirements, as well as specific annotations for guiding energy assessment [15].

In summary, one can understand that an environmentally sustainable (green) business process is a business process that generates business value with minimal impact on the environment, thus without compromising the environmental resource availability for future generations.

2.3 The Six Core Elements of Business Process Management

In this study, we consider the model of BPM capabilities [8] (Table 1) to analyze the literature on Green BPM in terms of its contributions related to the six core elements of BPM: strategic alignment, governance, methods, information technology, people, and culture. This model seems very appropriate for this study as all these elements represent critical success factors for BPM in an organization [8]. Therefore, each element should be considered by organizations seeking success with BPM [8]. For each element, distinct capability areas were identified, as shown in Table 1. The six core elements can be described as follows [8]:

- **Strategic Alignment**: The complete cycle of BPM needs to be tightly linked with the overall strategy of the organization, enabling continuous improvement and achievement of business goals.
- **Governance**: BPM Governance establishes pertinent and straightforward accountability for different levels of BPM. In addition, it provides support to the design of decision-making and compensation processes to manage process-related actions in business process management.
- **Methods**: Set of instruments and techniques that enable and support consistent actions along the BPM lifecycle.

- **Information Technology**: IT-based solutions are the software, hardware, and information systems that enable and support BPM activities.
- **People**: Individuals and groups who continuously develop and apply their BPM-related skills and knowledge in order to enhance business performance.
- **Culture**: Culture is the collective beliefs and values that define BPM-related attitude and behavior in order to enhance business performance.

Table 1. The six core elements of BPM [8]

Strategic Alignment	*Governance*	*Methods*	*Information Technology*	*People*	*Culture*
Process Improvement Planning	Process Management Decision Making	Process Design & Modeling	Process Design & Modeling	Process Skills & Expertise	Responsiveness to Process Change
Strategy & Process Capability Linkage	Process Roles and Responsibilities	Process Implementation & Execution	Process Implementation & Execution	Process Management Knowledge	Process Values & Beliefs
Enterprise Process Architecture	Process Metrics & Performance Linkage	Process Monitoring & Control	Process Monitoring & Control	Process Education	Process Attitudes & Behaviors
Process Measures	Process-Related Standards	Process Improvement & Innovation	Process Improvement & Innovation	Process Collaboration	Leadership Attention to Process
Process Customers & Stakeholders	Process Management Compliance	Process Program & Project Management	Process Program & Project Management	Process Management Leaders	Process Management Social Networks

3 Methodology

We conducted a structured literature review [16] to answer our research question presented in Section 1. The subject of this literature review is environmentally sustainable business process management. Therefore, the search terms were “Green”, “Sustain*”, “BPM”, “Business Process Management”, “Business Process”, and “Process”. *Google Scholar* was the main search engine used for the literature search, but *ScienceDirect* and *ABI/INFORM Complete* were also considered to corroborate the search process. The terms *Green IT* and *Green IS* were excluded from the search, since we aimed to find BPM-related papers exclusively. Initially, we retrieved 52 papers that we considered relevant for this study based on their title and abstract. However, we did not have access to full-text of ten papers, thus these ones were excluded for further analysis. The final set contained 42 articles selected for deeper analysis, as they were identified to be relevant to our study, and we had access to their full content.

Based on the framework described on the previous section, we classified the literature according to its main coverage content. For each source, the categorization took into account frameworks, models, and propositions related to Green BPM. A paper that, for instance, simply mentions that strategic alignment is important in BPM, or that Information Technology plays a critical role in delivering process innovation, would *not* be categorized in the Strategic Alignment and Information Technology capability areas, respectively.

4 Findings

The analysis of the literature led us to two main categories of papers: (i) papers that provide a more general view on Green Business Process Management (10), such as literature reviews, or papers presenting holistic approaches covering many aspects of business process management; and (ii) papers with a focus on one or more specific areas of interest for Green BPM (32). Additionally, 67% of the papers analyzed were conceptual in nature.

The papers classified as *General* present: (i) differences and commonalities between green and conventional business process management [2]; (ii) a literature review on Green BPM [5]; (iii) a literature review and research framework for Green BPM [6]; (iv) a literature review on Green BPM [7]; (v) an introduction to and a framework for Green BPM research and practice [11]; (vi) an approach that utilizes a BPM framework in order to transform and manage business processes for a Green Telecommunications company [17]; (vii) ways to manage sustainability in the face of unordered business processes [18]; (viii) opportunities and challenges of Green BPM based on conceptual considerations [19]; (ix) a Green BPM readiness model [20]; and (x) the role of business process management in green initiatives [21].

With regards to the core elements of BPM, this study shows that most of the papers (54%) are focused on *Methods* (See figure 1). *Methods* and *Information Technology* together represent more than 80% of the literature. Our analysis thus shows that Green BPM research has been mainly focusing on the process lifecycle, aiming to find answers related to issues in the design, measurement, and improvement of processes in terms of environmental sustainability.

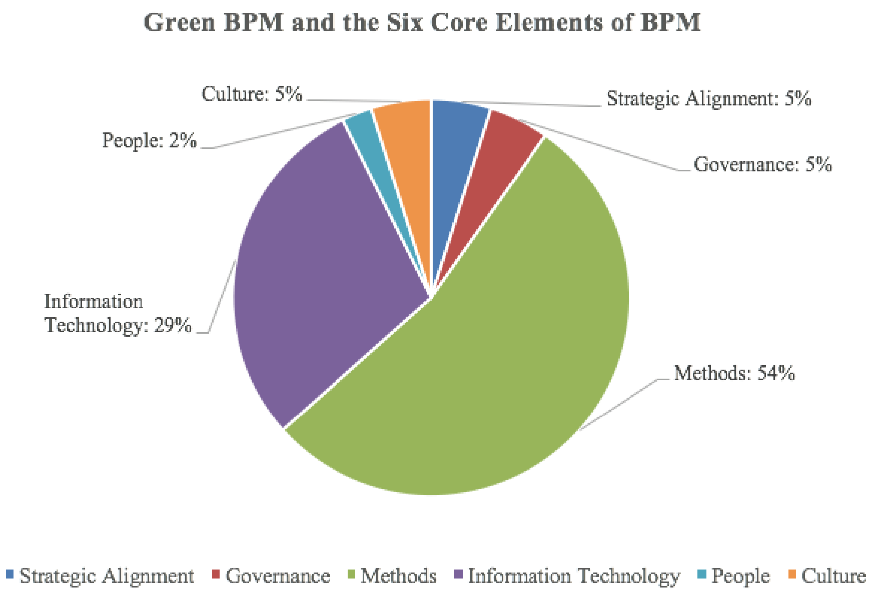


Figure 1. Green BPM Research on the lens of the six core elements of BPM

Table 2. (Green) Strategic Alignment

Topic	*Authors*
Business Motivation Model	[22]
Green Performance Indicators (GPIs)	[15]
Key Ecological Indicators	[2], [23], [24], [25]
Sustainability Indicators	[26], [27]
Process Architecture	[22]
Sustainability Balanced Scorecard (SBSC)	[28]

Table 3. (Green) Governance

Topic	*Authors*
Roles and Responsibilities	[2]
Process Performance Management	[29]
Process-related Standards	[30]

Table 4. (Green) Methods

Topic	*Authors*
Green Business Process Patterns	[12], [25], [31], [32]
Annotations	[15], [33]
Emission Annotations	[4], [34]
PMapping extension	[35]
Extensions of process modeling notations	[26], [36], [37]
Evaluation of process modeling notations	[38]
Activity-Based Emission (ABE)	[36], [39]
Green Activity Based Management (ABM)	[40]
Process Viewing Patterns	[24]
Enterprise Topology Graph (ETL)	[25]
Business Process Simulation	[26], [37]
Process SEER	[4], [34]
Abnoba Framework	[2], [5], [28], [41]

Table 5. (Green) Information Technology

Topic	*Authors*
Sustainability aware software system engineering framework	[27]
Evaluation of process modeling tools	[38]
Functional affordances of information systems	[14]

Topic	Authors
Energy Informatics	[42]
Process Mining	[13]
Geographic Information Systems (GIS)	[33]

Topic	Authors
New functionalities for process modeling and analysis tools	[34]
Process Simulation tools	[13]
Process Automation	[12]
IT solutions for Process Monitoring & Control	[25]
Energy-aware business processes	[15]
Conceptual integration model for the energy consumption of business processes, applications, and IT devices	[10]
Semantic Process Benchmarking	[43]
Cloud Patterns	[23]

Table 6. (Green) People

Topic	Authors
Collaborative Green Business Process Management	[44]
Process Education	[3]

Table 7. (Green) Culture

Topic	Authors
Responsiveness to process change	[45]
Process attitudes and behaviors for small and medium enterprises	[46]

5 Green BPM and the Process Lifecycle

Based on the literature review, we derived a **Green BPM lifecycle** (Figure 2 - adapted from [47]), which considers the methods, techniques, and approaches yielded through our literature analysis. The model represents a non-exhaustive rack of items found in the literature. This lifecycle model thus identifies which methods, techniques, and approaches might be selected and combined according to the specific business context, and under consideration of the environmental, process-related goals of an organization. Some of the methods present in the lifecycle (e.g. Process Viewing Patterns, process automation) are not exclusively designed for Green BPM. Nonetheless, we assume that, whether combined with other techniques or not, they might support the overall enterprise-wide process management framework, thus being able to contribute to making business processes more environmentally sustainable. We believe that the Green BPM lifecycle can serve as a framework for future studies in this field, and we thus encourage its empirical application and evaluation.

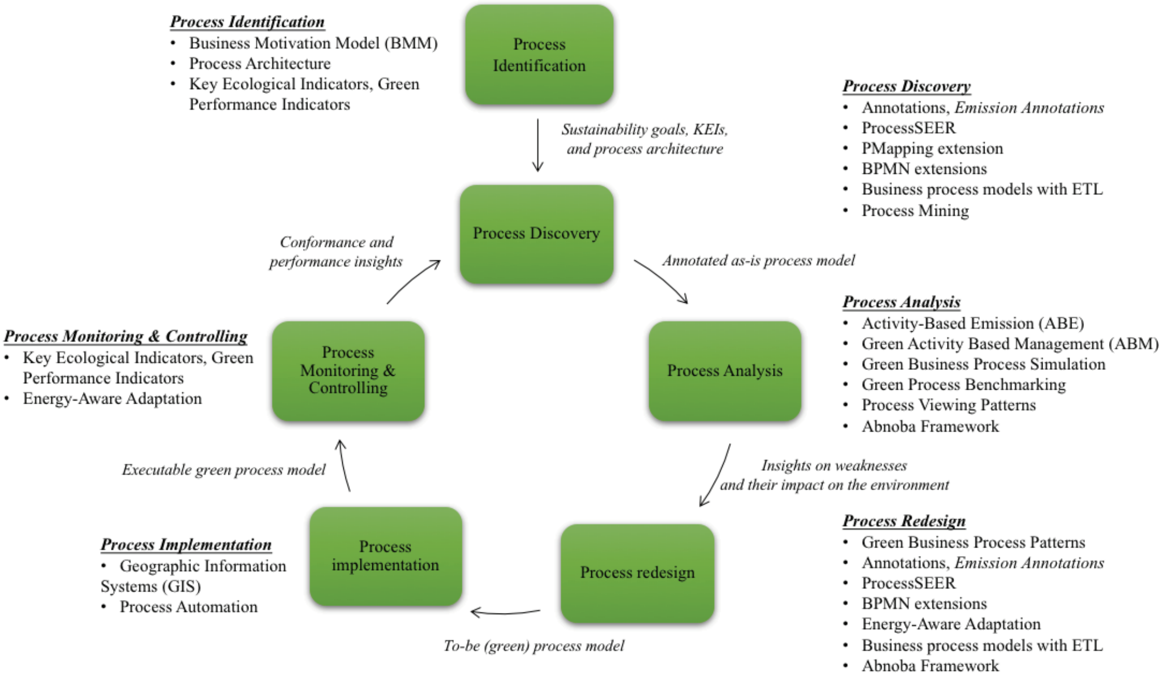


Figure 2. Green BPM lifecycle (Adapted from [47])

6 Research Agenda

From the literature analysis, we found several approaches that integrate sustainability-related thinking into business process management. Since most of the papers are conceptual, we believe that future research should focus on empirical studies that propose and evaluate methods, techniques, and initiatives that can be applied to all six core areas of BPM, and that can support organizations in the execution of business processes leading to as little impact as possible on the natural environment.

Strategic Alignment is mostly considered in terms of *Process Measures*, with environmentally sustainable performance indicators [2], [15], [23], [24], [25], [26], [27]. However, little or no attention has been paid to the other capability areas within this core component of BPM (e.g., process improvement planning, etc. – see Table 1). Future studies might focus on the alignment of Green BPM with the overall strategy of an organization. So, how could organizations address the trade-off between economic and environmental goals in their strategy? How to consider environmental characteristics of processes in the *Enterprise Process Architecture*? Which sustainability factors should be acknowledged in a *Process Improvement Plan*?

We found several **Methods** that can support the adoption of Green BPM and help organizations reduce the negative impact of their business processes. Nevertheless, there was no paper classified in the *Process Program & Project Management* capability area, and little attention has been paid to *Process Implementation & Execution*. So, how can sustainability factors be embedded into executable business process specifications? Which methods and approaches should be used for the management of green BPM projects?

Even though some relevant research on **Information Technology** as an enabler of Green BPM exists, there is still a lack of explicit focus on the process lifecycle. This is the case, for example, of process modeling tools that incorporate green process modeling notation extensions (e.g. BPMN, EPC) or BPM systems that consider ecological characteristics. How should these extensions be described in order to facilitate their implementation by BPM software vendors? Which tools could facilitate the management of green initiatives? We strongly encourage more in-depth research on solutions that support the mining, modeling, analysis, simulation, implementation, and monitoring of green business processes, thus providing means of integration and incorporation of sustainability factors within those information systems.

Our analysis indicates that research has paid little attention to **Governance, People,** and **Culture** in green business process management. We thus argue that there is still a need for studies on Green BPM that address these areas in more detail. In the context of *Governance,* what would be appropriate methods to measure an organization's Green BPM maturity? How to design structured and consistent decision-making processes to guide sustainability-related actions? When considering *People*, what are the *Process Skills and expertise* necessary in a sustainable BPM approach? For *Process Education* initiatives, what needs to be considered in a Green BPM education curriculum? In terms of *Culture* in Green BPM, what are the essential values and attitudes for meeting Green BPM objectives? What is the role of social networks in Green BPM?

7 Conclusion

This study aimed to analyze the current state of research on Green BPM. To that end, we conducted a literature review, which comprised several studies on Green Business Process Management that were categorized according to the six core elements of BPM. We found that most of the contributions were presented as conceptual studies and that the majority of papers were focused on methods and techniques that consider sustainability factors in business process management. We contributed with the description of the Green BPM lifecycle based on the methods, techniques, and approaches derived from the analysis of the literature. We argued that the Green BPM lifecycle can serve as a framework for future studies that aim to apply, combine, and compare methods that can be utilized in each phase of the business process management lifecycle to help organizations become more environmentally sustainable, thus reducing their impact on the environment. We discussed key gaps regarding Green BPM that emerged from our analysis, and suggested avenues for future research for each core element. We argued that future research should consider all six elements of BPM, and should evaluate and propose means to address the environmental impact of business processes. Furthermore, this study is not without limitations. First, we did not have access to full-text versions of all papers, so some relevant contributions might not have been included in this study. Having access to full-text papers for all literature found could have disclosed additional Green BPM approaches. Second, the paper coding was conducted by only one person, which might have led to biased paper classification. It might have been better if two researchers had been involved working on a pair coding mode. Still, by using the criteria provided by the BPM maturity model [8] the author took care to base the coding on formal concepts derived from the literature.

Acknowledgments

This paper has been written in the context of the Student Research Fellowship at the Institute of Information Systems at the University of Liechtenstein. The author is pleased to acknowledge Prof. Dr. Stefan Seidel for his support in the entire process. The author thanks the anonymous reviewers for their very constructive remarks.

References

1. Vom Brocke, J., Seidel, S., & Recker, J. (Eds.). (2012). Green Business Process Management: Towards the Sustainable Enterprise. Springer.
2. Nowak, A., Leymann, F., & Schumm, D. (2011). The Differences and Commonalities between Green and. In Proceedings of the International Conference on Cloud and Green (pp. 569-576). IEEE Computer Society.
3. Lan, Y. C. (2012). Reengineering a Green Business. International and Interdisciplinary Studies in Green Computing, 1.
4. Hoesch-Klohe, K., Ghose, A., & Lê, L. S. (2010, July). Towards green business process management. 2010 IEEE International Conference on Services Computing.
5. Gohar, S. R., & Indulska, M. (2015). Business process management: saving the planet?. In Australasian Conference on Information Systems (ACIS).
6. Opitz, N., Krüp, H., & Kolbe, L. M. (2014b, January). Green Business Process Management--A Definition and Research Framework. In System Sciences (HICSS), 2014 47th Hawaii International Conference on (pp. 3808-3817). IEEE.
7. Stolze, C., Semmler, G., & Thomas, O. (2012). Sustainability in Business Process Management Research–a Literature Review.
8. Rosemann, M., & vom Brocke, J. (2015). The six core elements of business process management. In Handbook on Business Process Management 1. Springer.
9. World Commission on Environment and Development (WCED) (1987) Our Common Future, Oxford University Press, UK and New York.
10. Reiter, M., Fettke, P., & Loos, P. (2014, January). Towards Green Business Process Management: Concept and Implementation of an Artifact to Reduce the Energy Consumption of Business Processes. In System Sciences (HICSS), 2014 47th Hawaii International Conference on (pp. 885-894). IEEE.
11. Seidel, S., Recker, J., & vom Brocke, J. (2012). Green business process management. In Green Business Process Management (pp. 3-13). Springer.

12. Nowak, A., Leymann, F., Schleicher, D., Schumm, D., & Wagner, S. (2011a, October). Green business process patterns. In Proceedings of the 18th conference on pattern languages of programs (p. 6). ACM.
13. Lubbecke, P., Reiter, M., Fettke, P., & Loos, P. (2015, January). Simulation-based Decision Support for the Reduction of the Energy Consumption of Complex Business Processes. 2015 48th Hawaii International Conference on (pp. 866-875). IEEE.
14. Seidel, S., & Recker, J. (2012, January). Implementing green business processes: the importance of functional affordances of information systems. Proceedings of the 23rd Australasian Conference on Information Systems 2012 (pp. 1-10). ACIS.
15. Cappiello, C., Fugini, M., Ferreira, A. M., Plebani, P., & Vitali, M. (2011, August). Business process co-design for energy-aware adaptation. 2011 IEEE 7th International Conference on Intelligent Computer Communication and Processing.
16. Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS quarterly, xiii-xxiii.
17. Balachandran, R. (2010). Business Processes Management for a Green Telecommunications Company. Handbook of Research on Green ICT: Technology, Business and Social Perspectives: Technology, Business and Social Perspectives, 197.
18. Hasan, H. (2012). Unordered business processes, sustainability and green IS. In Green Business Process Management (pp. 39-58). Springer.
19. Houy, C., Reiter, M., Fettke, P., & Loos, P. (2011, January). Towards Green BPM–Sustainability and resource efficiency through business process management. In business process management workshops (pp. 501-510). Springer.
20. Opitz, N., Krüp, H., & Kolbe, L. M. (2014a). Environmentally sustainable business process management–developing a green bpm readiness model.
21. Seidel, S., vom Brocke, J., & Recker, J. C. (2011). Call for action: investigating the role of business process management in green IS. Sprouts: Working Papers on Information Systems, 11(4).
22. Rozman, T., Draghici, A., & Riel, A. (2015). Achieving Sustainable Development by Integrating It into the Business Process Management System. In Systems, Software and Services Process Improvement (pp. 247-259). Springer.
23. Nowak, A., Leymann, F., & Mietzner, R. (2011c). Towards green business process reengineering. In Service-Oriented Computing (pp. 187-192). Springer.
24. Nowak, A., Leymann, F., Schumm, D., & Wetzstein, B. (2011d). An architecture and methodology for a four-phased approach to green business process reengineering. In Information and Communication on Technology for the Fight against Global Warming (pp. 150-164). Springer.
25. Nowak, A., Binz, T., Leymann, F., & Urbach, N. (2013, September). Determining Power Consumption of Business Processes and their Activities to Enable Green Business Process Reengineering. In Enterprise Distributed Object Computing Conference (EDOC), 2013 17th IEEE International (pp. 259-266). IEEE.
26. Betz, S. (2014). Sustainability aware Process Management using XML-Nets. In Proceeding of the 28th EnviroInfo Conference, Oldenburg.
27. Betz, S., & Caporale, T. (2014, December). Sustainable Software System Engineering. In Big Data and Cloud Computing (BdCloud), 2014 IEEE Fourth International Conference on (pp. 612-619). IEEE.
28. Houy, C., Reiter, M., Fettke, P., Loos, P., Hoesch-Klohe, K., & Ghose, A. (2012). Advancing business process technology for humanity: Opportunities and challenges of green BPM for sustainable business activities. In Green Business Process Management (pp. 75-92). Springer.

29. Cleven, A., Winter, R., & Wortmann, F. (2012). Managing process performance to enable corporate sustainability: a capability maturity model. In Green Business Process Management (pp. 111-129). Springer.
30. Lange, M. I., & PMP, P. (2013). Consideration of sustainable development principles in process management.
31. Nowak, A., Binz, T., Fehling, C., Kopp, O., Leymann, F., & Wagner, S. (2012). Pattern-driven green adaptation of process-based applications and their runtime infrastructure. Computing, 94(6), 463-487.
32. Nowak, A., & Leymann, F. (2013, December). Green Business Process Patterns--Part II (Short Paper). In 2013 IEEE 6th International Conference on Service-Oriented Computing and Applications (pp. 168-173). IEEE.
33. Zhu, X., Zhu, G., vanden Broucke, S., & Recker, J. (2015). On Merging Business Process Management and Geographic Information Systems: Modeling and Execution of Ecological Concerns in Processes. In Geo-Informatics in Resource Management and Sustainable Ecosystem (pp. 486-496). Springer.
34. Ghose, A., Hoesch-Klohe, K., Hinsche, L., & Le, L. S. (2010). Green business process management: A research agenda. Australasian Journal of Information Systems, 16(2).
35. RT White, G., & James, P. (2014). Extension of process mapping to identify "green waste". Benchmarking: An International Journal, 21(5), 835-850.
36. Recker, J., Rosemann, M., Hjalmarsson, A., & Lind, M. (2012). Modeling and analyzing the carbon footprint of business processes. In Green Business Process Management (pp. 93-109). Springer.
37. Wesumperuma, A., Ginige, J. A., Ginige, A., & Hol, A. (2011). A Framework for Multi-Dimensional Business Process Optimization for GHG Emission Mitigation.
38. Opitz, N., Erek, K., Langkau, T., Kolbe, L., & Zarnekow, R. (2012). Kick-starting Green Business Process Management–Suitable Modeling Languages and Key Processes for Green Performance Measurement.
39. Recker, J., Rosemann, M., & Gohar, E. R. (2011, January). Measuring the carbon footprint of business processes. In Business process management workshops (pp. 511-520). Springer.
40. Wesumperuma, A., Ginige, A., Ginige, J., & Hol, A. (2013). Green activity based management (ABM) for organisations. In 24th Australasian Conference on Information Systems (ACIS) (pp. 1-11). RMIT University.
41. Hoesch-Klohe, K., & Ghose, A. (2010, December). Carbon-aware business process design in Abnoba. In International Conference on Service-Oriented Computing (pp. 551-556). Springer.
42. Watson, R. T., Howells, J., & Boudreau, M. C. (2012). Energy Informatics: Initial thoughts on data and process management. In Green Business Process Management (pp. 147-159). Springer.
43. Gräuler, M., & Teuteberg, F. (2013). Experimental Evaluation of a Process Benchmarking Tool in a Green Business Process Management Context. In Wirtschaftsinformatik (p. 68).
44. Jakobi, T., Castelli, N., Nolte, A., Stevens, G., & Schönau, N. (2014). Towards Collaborative Green Business Process Management. In EnviroInfo (pp. 683-690).
45. Gautier, P., Fry, C., Fedrigo, C., Hild, P., & Takagi, A. (2012, January). How to assess a green process: the green ROI. In ISPIM Conference Proceedings (p. 1).
46. Baggia, A., Leskovar, R., Delibašić, B., & Petrović, N. Opportunities of sustainable business practices in SME's.
47. Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2013). Fundamentals of business process management (Vol. 1, p. 2). Springer.

Reconstructing the Giant: Automating the Categorization of Scientific Articles with Deep Learning Techniques

David Dann[1], Matthias Hauser[1], and Jannis Hanke[1]

[1] University of Würzburg, Faculty of Management and Economics, Würzburg, Germany
mail@daviddann.de, {matthias.hauser,jannis.hanke}@uni-wuerzburg.de

Abstract. The present paper is concerned with the automation of the process of conducting literature reviews. Manual reviews are getting more and more difficult as the number of publications increases steeply. Against this backdrop, we investigate the application of deep learning techniques for the automation of the time consuming step of comparing and categorizing large sets of scientific literature. In contrast to prior research, we leverage the potential of the word2vec algorithm that provides a more semantic focus of analysis than common text mining approaches. We evaluate our artifact considering an exemplary document collection comprising 906 articles on Radio Frequency Identification. Our results indicate that our word2vec-based system provides better results than a system based on traditional text mining approaches.

Keywords: Text Mining, Literature Review, Word2vec, Design Science, RFID

1 Introduction

Free text is the most natural form of storing information. Nair and Narayanan [1, 2] show that up to 80% of the world's data is stored in the form of unstructured data such as text documents and that such data is growing at 15 times the rate of structured data. In times of proceeding digitalization, capturing the content of this growing data pool becomes increasingly difficult. As a result, the means to structure and understand unstructured data become more and more important.

This problem is particularly evident in the process of creating literature reviews, which summarize the current state of research in a scientific field and are thus a fundamental component in the process of knowledge creation. The Thomson Reuter's Web of Science, for example, contains about 58 million articles. This large number is clearly illustrated by Van Noorden et al. [3] who outline that printing just one page of every item would lead to a stack of papers that would reach almost to the top of Mt. Kilimanjaro. Moreover, Bornmann and Mutz [4] show that the number of publications doubles approximately every 24 years. Obviously, the large number of articles make it difficult to identify relevant information [5]. We conclude that there is a need for improving the process of conducting literature reviews in order to successfully process this increasing number of articles.

13th International Conference on Wirtschaftsinformatik,
February 12-15, 2017, St. Gallen, Switzerland

Dann, D.; Hauser, M.; Hanke, J. (2017): Reconstructing the Giant: Automating the Categorization of Scientific Articles with Deep Learning Techniques, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1538-1549

To address this information overload problem, firstly, the literature review process should be structured and follow well established frameworks (e.g., [6–8]). This helps the creation of consistent and comparable reviews and allows researchers to extend easily the work of others and so keep the research community up to date [6]. Secondly, techniques for partial automation of the process should be investigated in order to keep up with the ever growing acceleration of article publications [5].

To facilitate the automation of reviews, a structured review process is a necessary prerequisite. In our research, we consider the structured process for conducting literature reviews described by vom Brocke et al. [6] and investigate the automation of the most labor-intensive process step. Their framework encompasses five phases. In the first phase, the focus of the review is defined. In the next step, one has to gain a broad conception of the selected research topic which covers, for example, the understanding of key concepts and the uncovering of relevant search terms. Afterwards, relevant articles for the review are identified based on these search terms. In the fourth step, the literature is analyzed and synthesized. Here, vom Brocke et al. [6] suggest using a concept matrix to categorize the literature. These matrices were adapted for literature reviews by Webster and Watson [7] and map individual articles to the concepts they belong to. In the last step of the framework, a research agenda is created based on the results from previous phases.

In our research, we aim at automating the fourth phase of the considered framework which is especially complex and time-consuming. To this end, we apply data mining techniques to compare and categorize semantic contents of large amounts of scientific articles. To approach this task, the text documents must first be transformed into a structured data form. For this step, systems described in literature mostly rely on the commonly known bag-of-words model (e.g., [5, 9–11]). This model, however, has several drawbacks and is less suited for categorizing literature owing to loss of word order and the linguistic phenomena of ambiguity and synonymy which are not well handled by the model [5]. In contrast, we develop a system based on the word2vec model introduced by Mikolov et al. [12]. This model uses high quality vector representations of words to express their semantic information value and therefore should be more suitable for comparing and categorizing literature.

2 Related Work

Algorithm-based comparisons of scientific documents and their potentials have been of great interest to researchers (e.g., [5, 9–11, 14–16]). These approaches can be roughly categorized by the way they transform unstructured data into structured data. We differentiate between (i) text-based, (ii) citation-based, and (iii) hybrid approaches. Text-based approaches compare the textual content of articles (e.g., [9–11]), link-based approaches the citation links of articles (e.g., [14, 16]), and finally, hybrid approaches build on both approaches (e.g., [5, 15]).

The text-based approaches we found in literature use the bag-of-words model for the data transformation step but consider different text sections (e.g., keywords, titles, abstracts, full texts) and pursue different objectives (e.g., categorization of documents or development of recommendation systems). Gulo et al. [10] analyze abstracts of articles and use machine learning techniques and Bayesian classifiers for categorizing these articles. In contrast, Wang and Blei [11] develop a recommendation system for scientific literature. To this end, they process abstracts and titles of scientific articles with topic modelling and collaborative filtering techniques. The document collection of Afonso and Duque [9] contains titles of articles, abstracts, keywords and the first page or column of the introductions. They again aim at categorizing scientific literature and therefore compare different automated text-based clustering approaches.

Carpenter and Narin [14] use citation-based approaches for the data transformation step. For this purpose, they assume that journals in the same scientific domain have similar reference patterns and thus refer primarily to each other. As a result, they base their clustering process on cross-citation links of articles. Chen [16] uses this assumption and develops a citation-based system for analyzing and structuring literature. Deploying principal component analysis, he generates a correlation value which serves as a measurement for the relatedness between scientific papers.

Hybrid approaches were proposed by Bolelli et al. [15] and Aljaber et al. [5]. As with the purely text based approaches, both author groups use the bag-of-words model for the data transformation step. However, in contrast to these approaches they also consider the citations of articles. The first group considers complete article texts and the citation graph spanned by the articles for the categorization of documents.[1] In contrast, the second group considers citation contexts which are sequences of words surrounding citation markers within the full texts of the documents. They argue that including the citation contexts addresses the issue of synonymy, one of the drawbacks of the bag-of-words model. This is because, according to the authors, these contexts provide "relevant synonymous and related vocabulary which will help increase the effectiveness of the bag-of-words representation". The authors furthermore benchmark their model against a purely citation-based and a text-based model that considers the full texts of documents. Their results indicate that citation-based approaches are inferior to text-based approaches for document categorizations.

Counter to Bolelli et al. [15] and Aljaber et al. [5], we choose a purely text-based approach. We do, however, not rely on the bag-of-words model for building word representations. Instead, we use the word2vec model introduced by Mikolov et al. [12], that is able to express the semantic information value of big amounts of data in a way the bag-of-words model is not. In contrast to the hybrid approach developed by Aljaber et al. [5], the word2vec model addresses not only the synonymy issue of the bag-of-words model but also the issue of ambiguity and considers the information value of the word order. This makes a combination of link-based and text-based approaches unnecessary and justifies our purely text-based approach.

[1] A citation graph is a graph with papers as vertices and citations as directed edges between citing and cited documents.

3 Design-oriented Research

We pursue a design-oriented research approach to develop an artifact that automatically compares and categorizes scientific literature. The artifact design process follows the guidelines put forward by Hevner et al. [13]:

- **Problem Relevance:** Conducting a structured literature review is a complex and time-consuming endeavor and is getting more and more difficult as the number of publications increases every day [4]. To cope with this trend, it is essential to examine whether available text and data mining techniques are suitable for automating the process of structuring scientific articles.
- **Research Rigor:** In our research, we utilize deep learning and established data mining techniques. The proposed artifact is based on the word2vec model introduced by Mikolov et al. [12] that is capable of processing large amounts of textual data.
- **Design as a Search Process:** The idea of conducting an automated literature review is not completely new, but has been a field of research for several years [9]. Our paper examines in particular whether the word2vec model is suited to automate the fourth phase of vom Brocke et al.'s [6] literature review framework.
- **Design as an Artifact:** We design an IT artifact consisting of three components to provide an automated categorization of scientific literature. The artifact is implemented using the Python programming language.
- **Design Evaluation:** We evaluate the artifact considering an exemplary document collection comprising 906 articles on Radio Frequency Identification. We benchmark the artifact against a system based on the bag-of-words model.
- **Research Contribution:** We propose an artifact that is able to process large numbers of scientific articles and conceptualize them. Therefore, we apply novel deep learning methods that provide a more semantically focus of analysis than ordinary text processing models.
- **Research Communication:** A system that allows to capture large amounts of scientific literature quickly and effectively is an enrichment for scientists from all research disciplines. In addition, our paper addresses an audience with a more technical focus by explaining in detail the design of the proposed artifact.

4 Artifact Description

The artifact architecture is depicted in Figure 1 and consists of three components. The first component generates a word vector model based on the user-generated document collection (Subsection 4.1). This model allows the representation of each document in a vector space. Subsequently, comparing these vectors enables determining the similarity of documents. The output of the first component is a matrix containing all the similarity values among the documents in the collection and is input to the artifact's second component. This component groups the documents based on this matrix with a hierarchical clustering algorithm (Subsection 3.2). Finally, in the third component each cluster is automatically labelled with meaningful keywords through the application of a keyword extraction algorithm (Subsection 4.3). The keywords describe the clusters and thus inform users about the predominant topics within the individual clusters.

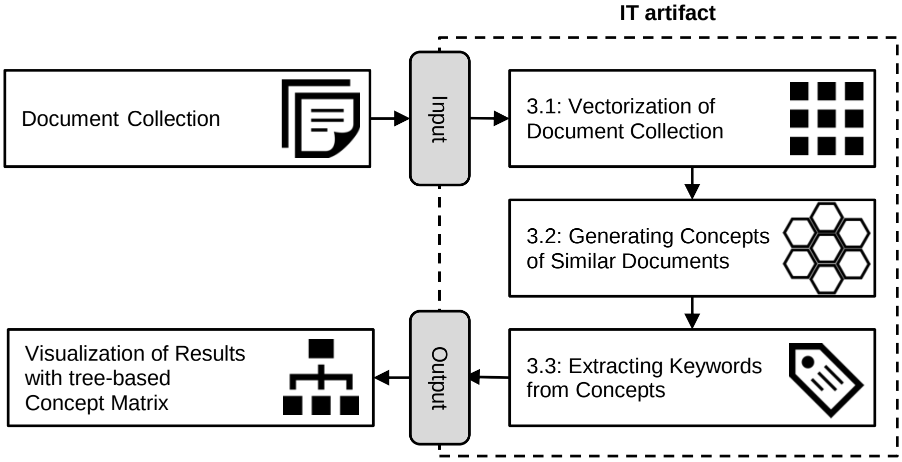


Figure 1. Architecture of the IT artifact

4.1 Vectorization of Document Collection

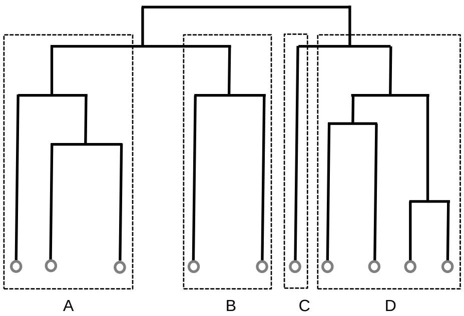
A user-generated collection of text documents serves as input for the proposed artifact. To represent the single documents as vectors, we train a data model using the Paragraph Vector algorithm introduced by Le and Mikolov [17] which is an extension of Mikolov et al.'s word2vec model [12].

Word2vec aims at training a word vector for each word occurring in the document collection using artificial neural networks. Using this deep learning technique enables us to prevent the drawbacks of the commonly used bag-of-words model. On the one hand, the bag-of-words model does not make use of the information value of the word order which may lead to errors because the model provides identical representations of semantically different sentences in case the same words are used [17]. On the other hand, the model cannot capture the linguistic phenomena of ambiguity and synonymy. The first phenomenon denotes lexically similar but semantically distinct words, the second semantically different but lexically similar words. These weaknesses lead to a more syntactic than semantic focus of the analysis, which makes the bag-of-words model less suitable for comparing and categorizing literature. In contrast, the word2vec model considers the context of words and thus eliminates the problems of traditional text processing models. Mikolov et al. [18] defines the word context as the words that surround a particular word. Leveraging these word contexts allows taking the word order into account. In addition, considering all individual word contexts in the entire document collection results in a high dimensional vector space in which vectors of semantically related words are located in close proximity to each another.

In order to make documents comparable, the Paragraph Vector algorithm builds on the word2vec model to represent each document as a concatenation of its word vectors in structured form. This vector representation of entire documents allows content-based similarity calculations using traditional distance measures. We apply the cosine distance, which is widely used in text mining applications to quantify the semantic relatedness of documents. We construct a document×document similarity matrix containing similarity values ranging from 0 to 1 with high values indicating similar contents and vice versa. The matrix is input to the artifact's second component.

4.2 Generating Concepts of Similar Documents

The second component automatically groups content-related documents based on the generated similarity matrix. We apply an agglomerative hierarchical clustering approach because it does not require an explicit specification of the number of clusters.[2] We employ the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) algorithm introduced by Sokal and Michener [19]. This algorithm iteratively compares all pairs of the assembled clusters based on the average distance of all elements within them. This allows the construction of a representation of all the documents in the collection in the form of a dendrogram (i.e., a tree-based diagram used for the visualization of clustering results). To receive homogeneous clusters, we automatically merge different branches of the generated dendrogram using the elbow-method introduced by Thorndike [20].[3] Each of the remaining clusters can be considered as columns in Webster and Watson's [7] concept matrix and thus be seen as a distinct concept (see Figure 2). Subsequently, to define the concepts, the artifact's third component applies keyword extraction techniques.



Articles	Concepts			
	A	B	C	D
1	✓			
2			✓	
3				✓
4	✓			
5		✓		
…				

Figure 2. Visualization of clustering results with dendrogram (left) and concept matrix (right)

[2] This approach builds clusters by initially treating each data point as its own cluster. The most similar clusters are then recursively merged until all data points form one single cluster.

[3] The elbow-method aims at finding the optimal number of clusters. The method, therefore, determines the clustering step in which merging two clusters leads to the maximum of variance between all clusters' centroids (which is equal to the maximum value of the second derivative of the average within-cluster distance values function).

4.3 Extracting Keywords from Concepts

The artifact's third component labels each concept with a predefined number of keywords that provide an impression of the documents' content. These keywords are directly drawn from the words occurring in the concept's original texts. To this end, we use the Rapid Automatic Keyword Extraction (RAKE) method developed by Rose et al. [21], which is an unsupervised, domain- and language-independent method for extracting keywords from text collections.[4] RAKE is particularly suited for our system due to its ability to pick highly specific terminology. As a result, our artifact generates meaningful concepts containing content-related scientific documents. As called for by vom Brocke et al. [6], the artifact thus provides a synthesis of scientific documents which facilitates working on the fifth phase of the framework – the generation of a research agenda.

5 Preliminary Evaluation

For evaluation, we instantiate our artifact based on an exemplary collection of scientific articles on Radio Frequency Identification. We consider articles containing the search term "RFID" in title, abstract or the full text.[5] The resulting document collection comprises 906 articles from 39 different journals and conference proceedings.

We benchmark our artifact against a system based on the bag-of-words model.[6] Before feeding the articles into the two systems, the titles, abstracts, keywords and reference sections were removed leaving only the articles' full texts. Based on the previously mentioned elbow-method, each of the systems identified 24 different concepts.

Figure 3 visualizes the systems' results. While the circular dendrogram generated with the bag-of-words-based system (left dendrogram) depicts one very large concept covering almost two thirds of the documents, the dendrogram generated with the word2vec-based system (right dendrogram) shows a more even distribution of concepts.

Figure 3 also zooms into the dendrogram generated with the word2vec model and lists the document titles of two exemplary concepts. Reading the titles of the documents in the concept displayed above the dendrograms suggests that all listed articles are

[4] The algorithm finds representative keywords by first splitting the text into text sequences using delimiters like punctuation or stop words. Then, for individual words in these sequences, word scores are calculated based on word frequency and word degree (i.e., the sum of the length of all sequences the particular word occurs in). The RAKE method then selects the top-scoring words as keywords.

[5] Following suggestions in literature (e.g., [6, 7]) we only consider high-quality articles. Therefore, we relied on the VHB-JOURQUAL 3 ranking [22] and examined articles in the sub-ratings "Operations Research" and "Wirtschaftsinformatik" that were at least B-ranked.

[6] We trained the word2vec model using the genism toolkit (http://radimrehurek.com/gensim/) and the bag-of-words model with the common term-frequency inverse document-frequency weighting using the scikit-learn library (http://scikit-learn.org/). In both cases we relied on the suggested default parameters.

healthcare-related. The titles listed below the dendrograms indicate articles about logistics and supply chain management. Given that the titles of the articles were not part of artifact's input, the results seem promising and indicate a comprehensible conceptualization.

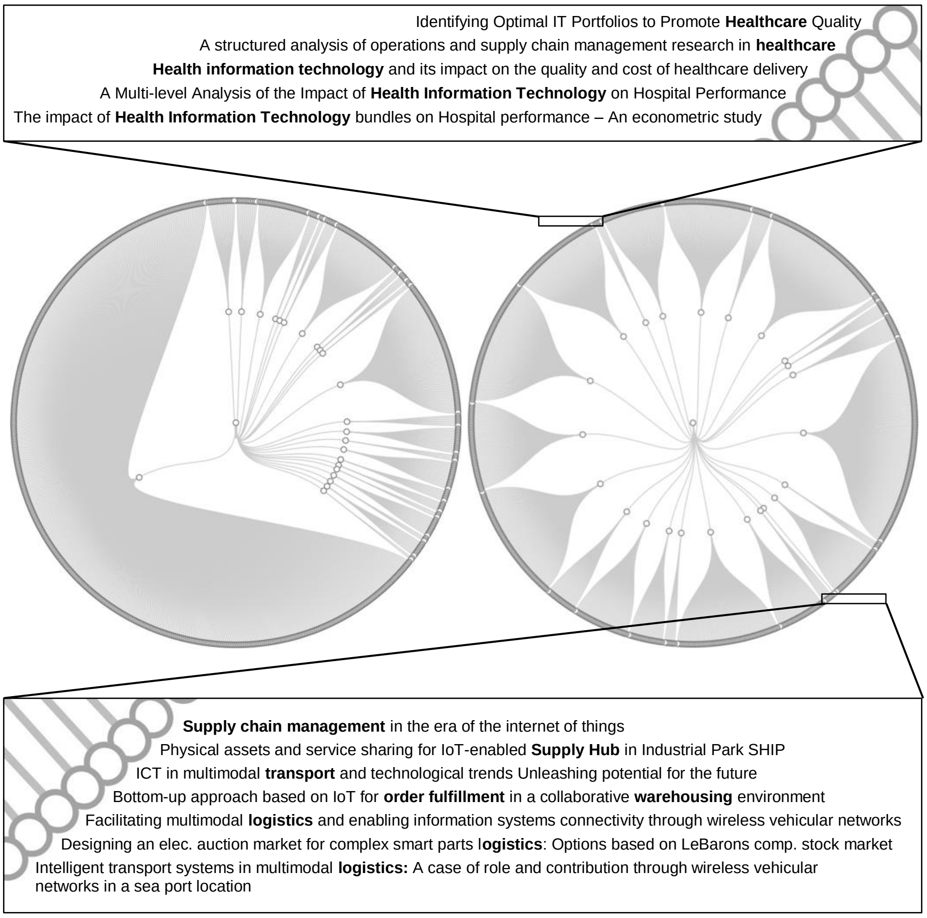


Figure 3. Circular dendrograms generated with bag-of-words-based system (left dendrogram) as well as word2vec-based system (right dendrogram) and two highlighted exemplary concepts

We consider two measures that allow a first quantitative assessment of the two systems' results. Our measures are based on assumptions similar to those introduced by Carpenter and Narin [14] who postulate that "journals which deal with the same subject area will have similar journal referencing patterns" and "journals which deal with the same subject area will refer to each other". Our underlying assumptions are:

1. The content of articles in a specific research domain (e.g., Operations Research) is more likely to be semantically related than content of articles in different research domains.
2. The content of articles in a specific journal (e.g., JAIS, EJOR) is more likely to be semantically related than content of articles in different journals.

We deduce from the first assumption that articles of a specific research domain should be more likely to get grouped into the same concept (research domain score). Second, we conclude from assumption 2 that articles of a specific journal should be more likely to get grouped into the same concept (journal score).

Therefore, we calculate the journal score by

$$\text{journal score} = \sqrt{\sum_{i}^{I}\sum_{j}^{J}(\frac{N_j}{N} - \frac{N_{ij}}{N_i})^2 \cdot N_j} \ ,$$

where I is the number of concepts, J the number of journals and conference proceedings, N_j the number of articles in j, N_i the number of articles in i, and N_{ij} the number of j's articles in i. The journal score compares the distribution of journals in individual concepts with the distribution of journals in the entire document collection.[7] An accumulation of one journal's articles within one concept thus leads to a high score.

In analogy, the research domain score is given by

$$\text{research domain score} = \sqrt{\sum_{i}^{I}\sum_{d}^{D}(\frac{N_d}{N} - \frac{N_{id}}{N_i})^2 \cdot N_d} \ ,$$

where D is the number of research domains. The research domain score can be considered as a generalization of the journal score. The higher the values of the score, the higher the accumulation of one domain's articles within one concept.

Table 1 summarizes the results for both systems. These preliminary results indicate that the bag-of-words-based system is inferior compared to the word2vec-based system. Although the results are very promising, further evaluations are necessary to fully evaluate the quality of the artifact.

Table 1. Evaluation Results

System	*Concepts*	*Journal score*	*Research domain score*
Bag-of-words-based	24	6.32	13.84
Word2vec-based	24	8.23	20.00

[7] Similar to the statistical measure standard deviation, which considers deviations of single observations, the journal score considers deviations of distributions within concepts. As the individual concepts differ in size, we normalize their squared deviations with their size.

6 Expected Contribution and Future Work

The understanding of an ever-growing number of scientific articles is an increasing challenge for any research domain. Literature reviews are an important instrument for structuring this information. They inform a research community about new findings and enable them to stay up-to-date. However, the number of available publications doubles approximately every 24 years [4]. This can make the process of structuring current research within a research domain very time-consuming, if not impossible. To cope with this trend, the present paper presents an IT artifact that leverages the potential of deep learning techniques to automate the time-consuming fourth phase of the literature framework proposed by vom Brocke et al. [6]. This phase comprises comparing and categorizing large amounts of potentially relevant literature.

Our research has a number of implications for both theory and practice. Regarding theory, we show that text- and data- mining techniques can be applied for this automation step and present an IT artifact capable of automatically categorizing large collections of scientific papers. Furthermore, our preliminary evaluation indicates that the clustering results benefit from the utilization of novel deep learning techniques. Due to the capabilities of Mikolov et al.'s [12] word2vec model to represent linguistic phenomena like ambiguity, synonyms or the information value of word order, it is better suited in the task of processing textual data than the common known bag-of-words model.

For practitioners (in this case scientific researchers), the study provides insights into the capabilities of an automated analysis of scientific document collections of arbitrary size. Implementations of automated literature categorization systems can (i) enable categorizing an amount of scientific literature which is impossible to handle manually, (ii) reduce the time needed to perform the complex but not complicated phase four of document analysis and synthesis and thereby (iii) improve both the researcher's productivity and the quality of the conducted literature review.

We consider the present paper as research in progress. As the evaluation of the artifact is a central activity in conducting rigorous Design Science Research [13], we are currently working on additional possibilities to demonstrate the artifact's capability. Firstly, we are working on including another score for assessing the quality of the results. To this end, we are implementing a measure considering citation patterns of scientific articles based on the assumptions proposed by Carpenter and Narin [14]. These authors identified related articles using a cross-citing matrix and a correlation measure to form article clusters referencing each other. Based on their assumption that articles with similar topics have a higher proportion of citations among themselves, these clusters will allow a direct comparison with the clusters generated by our proposed artifact. In addition, we aim at gathering empirical evidence for the validation of our system leveraging the focus group approach adapted for design science by Tremblay et al. [23]. This quantitative evaluation approach seems especially suitable because it enables direct interactions with domain experts and potential users of our artifact.

References

1. Nair, R., Narayanan, A.: Benefitting from big data: leveraging unstructured data capabilities for competitive advantage. Booz Co. 2, (2012).
2. Nair, R., Narayanan, A.: Getting Results from Big Data-a Capabilities-Driven Approach to the Strategic Use of Unstructured Information. Booz Co. (2012).
3. Van Noorden, R., Maher, B., Nuzzo, R.: The top 100 papers. Nature. 514, 550–553 (2014).
4. Bornmann, L., Mutz, R.: Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references: Growth Rates of Modern Science: A Bibliometric Analysis Based on the Number of Publications and Cited References. J. Assoc. Inf. Sci. Technol. 66, 2215–2222 (2015).
5. Aljaber, B., Stokes, N., Bailey, J., Pei, J.: Document clustering of scientific texts using citation contexts. Inf. Retr. 13, 101–131 (2010).
6. Brocke, J. vom, Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Reconstructing the giant: On the importance of rigour in documenting the literature search process. In: 17th European Conference on Information Systems, ECIS 2009, Verona, Italy, 2009. pp. 2206–2217 (2009).
7. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Q. 26, xiii–xxiii (2002).
8. Schryen, G., Wagner, G., Benlian, A.: Theory of knowledge for literature reviews: an epistemological model, taxonomy and empirical analysis of IS literature. In: International Conference on Information Systems (ICIS) (2015).
9. Afonso, A.R., Duque, C.G.: Automated Text Clustering of Newspaper and Scientific Texts in Brazilian Portuguese: Analysis and Comparison of Methods. J. Inf. Syst. Technol. Manag. 11, 415–436 (2014).
10. Gulo, C.A.S.J., Rúbio, T.R.P.M., Tabassum, S., Prado, S.G.D.: Mining Scientific Articles Powered by Machine Learning Techniques. In: 2015 Imperial College Computing Student Workshop (ICCSW 2015) (2015).
11. Wang, C., Blei, D.M.: Collaborative topic modeling for recommending scientific articles. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. p. 448. ACM Press (2011).
12. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient Estimation of Word Representations in Vector Space. CoRR. abs/1301.3781, (2013).
13. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. MIS Q. 28, 75–105 (2004).
14. Carpenter, M.P., Narin, F.: Clustering of scientific journals. J. Am. Soc. Inf. Sci. 24, 425–436 (1973).
15. Bolelli, L., Ertekin, S., Giles, C.L.: Clustering Scientific Literature Using Sparse Citation Graph Analysis. In: Fürnkranz, J., Scheffer, T., and Spiliopoulou, M. (eds.) Knowledge Discovery in Databases: PKDD 2006. pp. 30–41. Springer Berlin Heidelberg, Berlin, Heidelberg (2006).
16. Chen, T.T.: The development and empirical study of a literature review aiding system. Scientometrics. 92, 105–116 (2012).

17. Le, Q.V., Mikolov, T.: Distributed Representations of Sentences and Documents. CoRR. abs/1405.4053, (2014).
18. Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J.: Distributed Representations of Words and Phrases and their Compositionality. CoRR. abs/1310.4546, (2013).
19. Sokal, R.R., Michener, C.D.: A statistical method for evaluating systematic relationships. Univ. Kans. Sci. Bull. 28, 1409–1438 (1958).
20. Thorndike, R.L.: Who belongs in the family? Psychometrika. 18, 267–276 (1953).
21. Rose, S., Engel, D., Cramer, N., Cowley, W.: Automatic Keyword Extraction from Individual Documents. In: Berry, M.W. and Kogan, J. (eds.) Text Mining. pp. 1–20. John Wiley & Sons, Ltd, Chichester, UK (2010).
22. Verband der Hochschullehrer für Betriebswirtschaft e.V.: VHB-JOURQUAL3. http://vhbonline.org/en/service/jourqual/vhb-jourqual-3/ (2015).
23. Tremblay, M.C., Hevner, A.R., Berndt, D.J.: Focus groups for artifact refinement and evaluation in design research. Commun. Assoc. Inf. Syst. 26, (2010).

Design Prinzipien für Microlearning Crowdsourcing-Systeme

Konzept für audiovisuelle Mediengestaltung

Damian Kogga [1], Nicole Krawietz [1], Fatih Cevik [1], Sven Brandau [1], Mahei Manhai Li [2]

[1] University of Kassel, Kassel, Germany
damianKogga@web.de, nicolekrawietz@arcor.de, fatih-cevik@gmx.de, sven.brandau@web.de

[2] University of Kassel, Information Systems, Kassel, Germany
mahei.li@uni-kassel.de

Abstract. Die vorliegende Arbeit befasst sich mit der Entwicklung von Designanforderungen und Designprinzipien und Gestaltungsrichtlinien für die audiovisuelle Kommunikation auf Crowdsourcing-Plattformen. Zur Identifizierung bestehender Probleme in der Praxis, wurde eine Prozessanalyse der Kommunikation auf einer Crowdsourcing-Plattformen analysiert, sowie durch Nutzer-Interviews Probleme bei der Kommunikation ermittelt. Um das identifizierte Problem zu verstehen, wurden mit Hilfe der Cognitive Load Theory drei Designanforderungen ermittelt. Daraufhin wurden nach den Prinzipien der Cognitive Theory of Multimedia Learning die Design Prinzipien und die Gestaltungsrichtlinien zur audiovisuellen Kommunikation entwickelt. Basierend auf den Erkenntnissen dieser Arbeit, kann eine weitere wissenschaftliche Beschäftigung mit computervermittelter Kommunikation, speziell auf Crowdsourcing Plattformen, und der Evaluation der vorliegenden Ergebnisse erfolgen.

Keywords: *Crowdsourcing, cognitive load, Design Science Research, Kommunikation*

1 Einleitung

Lernen durch Selbsthilfe ist eine oft gewählte Form der Problemlösung. Vor allem IT-basierte Systeme werden als hilfreiche Tools zur Übertragung von Wissen in Unternehmen betrachtet [1]. Durch die effiziente Nutzung der eigenen Kompetenzen

13th International Conference on Wirtschaftsinformatik,
February 12-15,2017, St. Gallen, Switzerland

Kogga, D.; Krawietz, N.; Cevik, F.; Brandau, S.; Li, M.M. (2017): Design Prinzipien für Microlearning Crowdsourcing-Systeme - Konzept für audiovisuelle Mediengestaltung, in Leimeister, J.M.; Brenner, W. (Hrsg.): Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), St. Gallen, S. 1550-1561

im Unternehmen können Wettbewerbsvorteile generiert werden [2]. Eine Art IT-gestützter Problemlösungsplattformen sind Crowdsourcing-Plattformen. Diese bieten Platz, auftretende Probleme mit anderen Nutzern zu teilen und Hilfestellungen von diesen zu erhalten. Online existieren bereits seit längerer Zeit allgemeine Plattformen wie Quora oder programmierspezifische Plattformen wie Stackoverflow.com. Diese existieren ebenfalls im unternehmensinternen Rahmen und sind in der Lösungsfindung für die Mitarbeiter sowie für das Unternehmen von Vorteil [3]. Für das Lösen von Problemen ist es von Bedeutung, die Mitarbeiter zur aktiven Mitarbeit zu motivieren und ihnen eine Plattform und Werkzeuge zur Hand zu geben, um möglichst gute Lösungsvorschläge zu erbringen. Aktuell beschränkt sich die Kommunikation von Microlearning Crowdsourcing-Plattformen, auf das Medium von Freitexten. Um eine weitere Möglichkeit zum Kommunizieren auf Crowdsourcing-Plattformen anzubieten, wird die Kommunikation per Videobotschaften eingeführt.

Etablierte Plattformen, wie z.B. Youtube oder Snapchat, bieten die Option der Kommunikation per Videos bereits mit großem Erfolg an. Durch dieses Alleinstellungsmerkmal grenzen sich diese Plattformen von der Konkurrenz ab und bieten ihren Nutzern einen Mehrwert. Dieser Mehrwert soll auf Crowdsourcing-Plattformen transferiert werden, was ein stärkeres sozio-emotionales Erleben ermöglicht und somit die Mitarbeiter zur Kommunikation auf Crowdsourcing-Plattformen motiviert [4]. Eine effiziente Videobotschaft zu erstellen und an andere weiterzugeben, ist für viele Mitarbeiter mit hohem Aufwand verbunden. Wichtige Anforderungen, wie eine angemessene kognitive Belastung, sollten bei dem Erstellen eines Videos beachtet werden. Durch Gestaltungsrichtlinien für die Videokommunikation können Mitarbeiter an die effiziente Erstellung eines Videos herangeführt und dabei unterstützt werden.

Dieses Paper beschäftigt sich gezielt mit der Frage, wie eine Richtlinie zum Generieren von Videos erstellt werden kann, um die Kommunikation auf Crowdsourcing-Plattformen effektiver zu gestalten. Dazu wurde in einem ersten Schritt der Kommunikationsprozess einer Crowdsourcing-Plattform näher betrachtet, um die Ist-Situation zu analysieren. Anschließend wurden anhand von semistrukturieren Interviews mit potentiellen Nutzern Probleme ermittelt, die bei der Kommunikation auf einer Crowdsourcing-Plattform auftreten. Basierend auf den neuen Kenntnissen wurden im letzten Schritt Design Anforderungen und Design Prinzipien aus der Cognitive Load Theory sowie der Cognitive Theory of Multimedia abgeleitet, welche für die Richtlinien zur Erstellung von Videos definiert wurden.

2 Methode

Der gestaltungsorientierte Forschungsansatz, auch bekannt als „Design Science Research" (DSR), wird als eine angewandte Form der Wissenschaft bezeichnet. Orientiert man sich an dem DSR Cycle von Vaishnavi und Kuechler [2004], so erhält man eine konkrete Vorgehensweise zur Durchführung eines gestaltungsorientierten Forschungsvorhabens [5]. Dieser idealtypische Forschungsprozess gestaltet sich als ein iteratives Modell, bestehend aus mehreren Phasen: Awareness of Problem, Suggestion, Development, Evaluation und Conclusion. Die vorliegende Arbeit orientiert sich partiell an Meth et al. [2015], indem die Herleitung von Design Anforderungen sowohl aus der Praxis, als auch theoriebasiert stattfindet und die Design Prinzipien (Solution Space) aus weiterer Literatur abgeleitet werden [6].

Da sich diese Arbeit mit dem Nutzungsverhalten auf Crowdsourcing-Plattformen beschäftigt, wurde in einem ersten Schritt der Nutzungsprozess auf einer Plattform betrachtet. Hierfür wurde auf eine Plattform eines führenden deutschen Crowdsourcingsanbieters zugegriffen. Der Prozess wurde anschließend in bestehende Erkenntnisse zu Crowdsourcing-Prozessen eingebettet und als Gesprächsgrundlage für die semistrukturierten Interviews herangezogen. Die Interviews wurden mit 12 potentiellen Nutzern im Alter von 24-52 Jahren durchgeführt. Die zentrale Frage bestand darin zu verstehen, wie man die Effizienz der Plattform steigern könnte. Die Untersuchungen zeigten, dass die Nutzer die Kommunikation als wenig ansprechend und, aufgrund des hohen Leseaufwands, teilweise auch anstrengend, empfanden. Nach dem Identifizieren der Probleme folgte eine Recherche der bestehenden Literatur [7], durch welche die Probleme mit Hilfe bestehender Theorien eingeordnet werden konnten. Dabei haben sich die Cognitive Load Theory nach Sweller [2011] sowie die Cognitive Theory of Multimedia Learning nach Mayer [2008] als geeignete Herangehensweisen für die Entwicklung von Richtlinien als Vorschlag zur Reduzierung der Probleme herausgestellt. Ziel war es, den Nutzern eine Lösung anzubieten, welche die von ihnen geäußerten Probleme adressiert und gleichzeitig die Potentiale in der Kommunikation ausschöpft. Die Basis der Herangehensweise bilden die drei Arten der kognitiven Belastung nach Sweller und seiner Annahme von den begrenzten Kapazitäten des Arbeitsgedächtnisses [8]. Diese wurden den Prinzipien nach Mayer [2008] zugeordnet.

3 Grundlagen

3.1 Crowdsourcing

Der Begriff des Crowdsourcing ist auf Howe [2010] zurückzuführen. Er beschreibt die Thematik als die Auslagerung von unternehmensinternen Aufgaben per offenen Aufruf an eine Gruppe von Menschen, der sogenannten Crowd [9]. Crowdsourcing lässt sich in zwei verschiedene Typen unterscheiden, dem internen und externen Crowdsourcing [10]. Da diese Arbeit sich mit Crowdsourcing-Nutzungsprozessen innerhalb von Unternehmen beschäftigt, wird sich ausschließlich auf das interne Crowdsourcing fokussiert. Als interne Crowd werden hier ausschließlich die Mitarbeiter eines Unternehmens verstanden, welche IT-gestützt miteinander kommunizieren [11]. Wir betrachten das Crowsourcingsystem, welches sich aus den Akteuren Crowdsourcers, Crowdsourcees und der Plattform zusammenstellt. Es verfolgt das Ziel der Selbsthilfe durch Microlearning.

3.2 Microlearning

Unter Microlearning versteht man eine Form des Lernprozesses, welche sich durch spezifische Eigenschaften auszeichnet. Das Besondere hierbei ist, dass das Lernen durch zeitliche kurze Lerneinheiten (auch als Learn Nuggets bekannt) definiert ist. Des Weiteren spielt bei dem Lernprozess die Komplexität des Inhalts eine elementare Rolle. Die Inhalte innerhalb der einzelnen Einheiten sollten einfach und leicht verständlich sein. Die einzelnen Schritte des Microlearnings sollten in ein umfassendes Gesamtkonzept eingebunden und durch ein passendes Medium vermittelt werden. Hierbei ist die Form des Mediums, wie beispielsweise über die Crowdsourcing-Plattform (Technologie), von besonderer Bedeutung [12]. Der Lernprozess des Microlearnings, der sich aus dem Microcontents zusammensetzt, kann von wenigen Sekunden bis zu 15 Minuten dauern [13]. An das Microlearning werden besondere didaktische Anforderungen, wie z.B. ein motivierend gestaltetes Feedback, gestellt [14].

4 Awareness of Problem

4.1 Ist-Zustand und Einbettung

Um auftretende Probleme bei der Nutzung von Crowdsourcing-Plattformen zu identifizieren, ist es zunächst wichtig, den Ist-Zustand zu analysieren. Dies geschah in Kooperation mit einem realen Crowdsourcing-Plattformanbieter. Dabei wurde der Prozess aus Abbildung 1 identifiziert.

Crowdsourcing-Plattformen zielen darauf ab, reale Probleme von Nutzern zu lösen [15]. Der gesamte Nutzungsprozess des Crowdsourcingsystems (CSS) teilt sich in drei Pools auf, in Crowdsourcers (Nutzer, die ein Problem haben und dieses als Frage formulieren) und Crowdsourcees (eine Gruppe von Nutzern, die auf Problemanfragen antworten) [16]. Bestimmte Crowdsourcees haben zudem eine besondere Moderatorenrolle, da diese aus der Masse an Lösungsvorschlägen durch Selektion eine Endlösung formulieren können. Der Prozess läuft wie folgt ab: Bei einem auftretenden Problem wird dieses im ersten Schritt vom Crowdsourcer auf einer Plattform gepostet. Anschließend wird die Problemstellung für alle Crowdsourcees sichtbar. Einzelne Nutzer haben dann die Möglichkeit, Antworten oder Lösungsvorschläge für die Problemstellung zu erstellen und zu veröffentlichen. Diese einzelnen Lösungsmöglichkeiten werden wiederum durch die Plattform für alle Crowdsourcess bereitgestellt und eine gemeinsame Diskussion bis zur endgültigen Lösungsfindung durch den Moderator erfolgt. Diese wird in einem letzten Schritt veröffentlicht und der Crowdsourcer erhält somit eine Antwort auf seiner Frage. Damit ist der Prozess beendet.

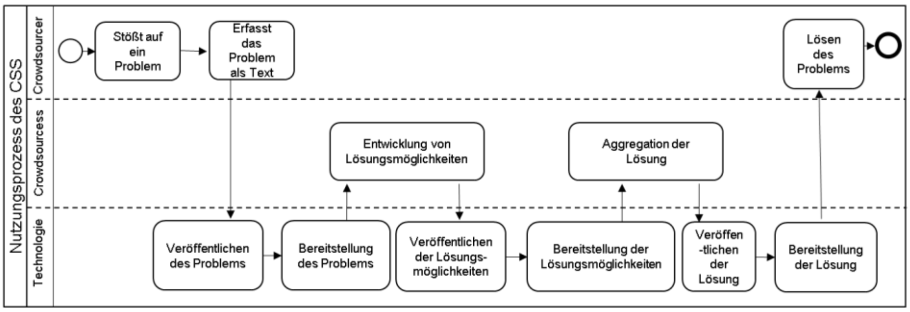


Abbildung 1: Nutzungsprozess einer internen Crowdsourcing Plattform *(Eigene Darstellung)*

Geiger et al. [2011] charakterisieren den Crowdsourcing-Prozess anhand von vier Phasen: 1. Preselection of contributors, 2. Accessibility of peer contributions, 3. Aggregation of contributions und 4. Remuneration for contributions [17]. In Phase 1 werden die Crowdmitglieder ausgewählt und in Phase 2 werden die jeweiligen Inhalte, die ein Crowdmitglied von einem anderen Crowdmitglied sehen kann, bestimmt. Phase 3 stellt das Zusammentragen von Lösungen dar. Phase 4 definiert, wie die Crowd für ihre Leistung entlohnt wird. Da in dieser Arbeit ausschließlich eine Betrachtung des internen Crowdsourcing im Unternehmen stattfindet, umfasst die Crowd alle Mitarbeiter des Unternehmens. Für diese sind alle Problemfragen offen sichtbar, sodass Phase 1, 2 und 4 ausgeklammert werden können. In Phase 3 findet die Kommunikation der Crowd durch das Teilen von erstellten Lösungsvorschlägen statt. Somit ist der Prozess aus Abbildung 1 primär in dieser Phase anzusiedeln und geht darüber hinaus, indem er sich explizit mit der Erstellung einzelner Beiträge beschäftigt.

4.2 Ergebnisse aus den Interviews

Die Aussagen der Befragten in den Interviews zeigten diverse Problembereiche auf. Diese lassen sich im Wesentlichen vier Kategorien zuordnen: Kommunikation, Motivationsbereitschaft zur Nutzung, Nutzerfreundlichkeit sowie allgemeine Bekanntheit der Plattform/des Plattformanbieters. Im Rahmen dieser Untersuchung erfolgt eine Fokussierung auf den Bereich der Kommunikation. Die Ergebnisse zeigten auf, dass acht der befragten Nutzer unzufrieden mit den bereitgestellten Kommunikationswerkzeugen sind und zudem fünf Nutzer den Wunsch nach visueller Kommunikation äußerten. Acht Befragte gaben an, dass sie die Kommunikation in alleiniger Textform als „zeitaufwändig" und zu „überwältigend" empfanden und es wurde fünf Mal der Wunsch nach Vereinfachung der Kommunikationsmöglichkeiten geäußert. Die Befragten empfanden das Verfassen in Textform als „langweilig und mühselig". Außerdem wurde die Annahme der audiovisuellen Videolösung als Darstellungsmedium durch den Vergleich mit Youtube-Tutorials durch die Befragten bestätigt. Folglich zeigen die Ergebnisse der Befragung den Wunsch nach der Erweiterung des Nutzungsprozesses in den Schritten „Erfasst das Problem als Text" und „Lösen des Problems" durch eine audiovisuelle Komponente.

4.3 Herleitung der Design Anforderungen

Zum besseren Verständnis und zur Diagnose der Gründe des Problems erfolgt auf Basis des „Activity Framework for Design Science Research", eine Auseinandersetzung mit dem Problemraum [18]. Zuerst erfolgt eine Ableitung von Design Anforderungen, sodass im nächsten Schritt die Problematik der Kommunikation effektiver gestaltet werden kann. Auf Basis dessen, sollen dann Design Prinzipien abgeleitet werden, welche die Nutzer der Crowdsourcing-Plattformen bei der Erstellung von Videobotschaften unterstützen. Es wird ein Augenmerk auf die didaktischen Anforderungen aus dem Bereich des Microlearnings gelegt. Die Lerninhalte, welche durch Videos kommuniziert werden, sollen möglichst verständlich, klar, kurz und effizient sein, um die Kommunikation und den Lernerfolg des Crowdsourcingsystems zu fördern [12, 13, 14]. Um dies zu gewährleisten werden Erkenntnisse aus der Cognitive Load Theorie herangezogen [8, 19, 20, 21]. Bei der Cognitive Load Theory nach Sweller [1988], werden zwei Annahmen getroffen: die Begrenztheit des Arbeitsgedächtnisses, was die These des Microlearnings unterstützt und dass das Arbeitsgedächtnis ein auditives und ein visuelles Zentrum für die Verarbeitung von Informationen besitzt. Um den Kommunikationsprozess effizient zu gestalten, ist es daher notwendig, die Informationen sowohl bildhaft als auch verbal zu übermitteln. Dieser Effekt wird auch als „Modality Effect" bezeichnet [8, 22]. Um die Anforderungen an die Design Prinzipien zu gestalten, werden im Folgenden die Arten der extrinsischen Belastung definiert und den in den Interviews identifizierten Problemen zugeordnet.

Festgestellt von den Nutzern wurde ebenfalls das Problem „Große Textmengen". Diese gehören nach Sweller der Art der extrinsisch kognitiven Belastung an [22, 23]. Viel Text führt dazu, dass die Nutzer konzentrierter arbeiten müssen und Ressourcen des Arbeitsgedächtnisses zusätzlich beansprucht werden. Da durch den Text nur das visuelle und nicht das auditive Zentrum angesprochen wird, findet ebenfalls eine Überlastung des visuellen Kanals statt [22, 23]. Als zweites Problem wurden die „schwierig zu verstehende Inhalte" genannt. Der Lerninhalt lässt sich der intrinsischen kognitiven Belastung zuordnen. Abhängig von der Komplexität des Lern/-Informationsinhaltes variiert der Belastungsgrad beim Nutzer [22, 24]. Als drittes Problem wurden die „schwierig zu vermittelnden Inhalte" genannt. Das dritte Problem kann teilweise der extrinsischen Belastung zuordnet werden, weil die schwierige Vermittlung des Inhaltes von der Gestaltung des Lernmaterials abhängig ist. Jedoch ist der schwierig zu vermittelnde Inhalt primär mit der lernbezogenen kognitiven Belastung erklärbar. Diese wird als der Aufwand verstanden, den Nutzer erbringen müssen, um das Lernmaterial zu begreifen [22]. Diese lernbezogene Belastung muss gefördert werden, um einen effizienten Informationstransfer zu gewährleisten. Die Bildung und anschließende Nutzung von Schemata hält die extrinsische kognitive Belastung gering und unterstützt bei der Aufnahme neuer Informationen [22].

Ziel der Cognitive Load Theory und somit eine wichtige Anforderung an die Design Prinzipien ist es, die kognitive Belastung möglichst gering zu halten, um effektiv Informationen aufnehmen bzw. kommunizieren zu können [8]. In Wissensübergabe umformuliert, wurden anhand der identifizierten Probleme folgende drei Design Anforderungen formuliert:

- *DR1. Für eine effektive Wissensübergabe durch Crowdsourcing-System entwickelten Lösungen, soll die extrinsische Belastung möglichst reduzieren.*
- *DR2. Für eine effektive Wissensübergabe durch Crowdsourcing-System entwickelten Lösungen, soll die intrinsische Belastung möglichst reduzieren.*
- *DR3. Für eine effektive Wissensübergabe durch Crowdsourcing-System entwickelten Lösungen, soll die lernbezogene Belastung möglichst erhöht werden.*

Um die in der qualitativen Befragung identifizierten Probleme zu verstehen, wurde in diesem Abschnitt die Cognitive Load Theory als theoretische Linse herangezogen. Daraus wurden die Design Anforderungen 1 – 3 abgeleitet, welche sich auf die Reduktion der kognitiven Belastung und somit die Entlastung des Arbeitsgedächtnisses beziehen.

5 Suggestion und Development

In diesem Schritt gilt es die Design Prinzipien zu entwickeln, welche einen effektiven Informationstransfer auf Crowdsourced-Microlearning-Plattformen ermöglichen. Hierfür werden die Arten kognitiver Belastung [8] und die bei den Interviews identifizierten Probleme den Prinzipien zugeordnet [18]. Ziel ist es, die Arten der kognitiven Belastung zu reduzieren, um das Arbeitsgedächtnis zu entlasten und einen effizienten Wissenstransfer zu ermöglichen. Auf dieser Basis werden Handlungsempfehlungen entwickelt, die die Nutzer bei der audiovisuellen Kommunikation unterstützen [20].

5.1 Herleitung der Design Prinzipien

Dass audiovisuelle Mitteilungen eine geeignete Maßnahme sind, die Kommunikation effektiv zu gestalten, wurde bereits von Paivio in seiner dual codingtheory festgestellt. Diese Theorie unterstützt Swellers und Mayers Annahme vom „Modality Effekt“ und besagt, dass das menschliche Gedächtnis unterschiedliche Repräsentationen besitzt, die von visueller sowie verbaler Natur sind [21, 22].

Auf diesen Erkenntnissen beruhend, entwickelt Mayer ein duales Model, welches sich auf die visuelle und auditive Informationsverarbeitung fokussiert und leitet daraus unterschiedliche Prinzipien ab [18]. Diese wirken sich förderlich auf den Wissenstransfer im multimedialen Lernen aus. Da die Erstellung von Microcontents als Lösungsvorschläge auch als Wissenstransfer betrachtet werden kann, wurden die als passend für den vorliegenden Betrachtungsgegenstand anerkannt. Die ersten fünf dieser Prinzipien beschäftigen sich mit der Reduktion extrinsisch kognitiver Belastung, die durch eine optimierte Verarbeitung relevanter Informationen stattfindet [19, 22].

Diese Prinzipien sind die Kohärenz (Reduktion von irrelevanten Informationen), die Signalisierung (Hervorhebung relevanter Informationen), die Redundanz (Vermeidung von Text bei Audioanimationen), die räumliche Kontiguität (Platzierung von Textzeilen neben der zugehörigen Grafik) und die temporale Kontiguität (Parallele Präsentation von visuellen und den dazugehörigen auditiven Inhalten) [8, 19]. Des Weiteren stellt Mayer Handlungsempfehlungen vor, die einen Umgang mit komplexen Sachverhalten ermöglichen. Diese sind die Segmentierung (das Aufteilen von Animationen in kleinere Segmente), das Pretraining (Trainings/ Schulungen über den zu behandelten Sachverlauf zum Beispiel Fachwörter), sowie die Modalität (Kombination visueller und auditiver Komponenten). Das Pretraining und die Modalität werden beim weiteren Vorgehen nicht berücksichtigt, da das Pretraining nicht vom Nutzer des CSS durchgeführt wird. Die Modalität wird ebenfalls nicht weiter beachtet, da dieses Prinzip bereits durch die Anwendung der Plattform auf Basis audiovisueller Kommunikation erfüllt wird [19].

Die letzten zwei Prinzipien wecken das Interesse des Nutzers an den zu vermittelnden Informationen. Dazu gehören Multimedia (Parallele Präsentation von Wörtern und Bildern ist besser als nur von Wörtern) sowie die Personalisierung (Informelle Informationsvermittlung wirkt sich positiver auf den Wissenstransfer als formelle aus). Da multimediale Quellen allein durch die Anwendung der Plattform erfüllt werden, werden diese ausgeklammert [19]. Die identifizierten Prinzipien liefern mögliche Lösungsansätze, um das Microlearning auf Crowdsourcing-Plattformen und somit die Kommunikation zu optimieren. Da der Wissenstransfer Microlearning CSS auf das multimediale Lernen abzielt, bilden die Prinzipien nach Mayer eine gute Basis für das weitere Vorgehen [8, 19, 20].

5.2 Aggregation von Design Anforderungen und Design Prinzipien

Wie das Zusammenspiel der aus der Theorie hergeleiteten Design Anforderungen, Design Prinzipien und konkreten Gestaltungsrichtlinien auf die aus der Praxis identifizierten Probleme ausüben, wird anhand Tabelle 1 erläutert:

Zu Beginn wurde das Problem „Große Textmenge“ durch die Interviews identifiziert. Dabei handelt es sich um eine extrinsische Belastung (DR1). Viel Text führt dazu, dass die Nutzer konzentrierter arbeiten müssen und Ressourcen des Arbeitsgedächtnisses zusätzlich beansprucht werden [8]. Um die kognitive Belastung zu reduzieren, sollte hier das Kohärenz Prinzip herangezogen werden. Irrelevante Informationen führen zu einer Erhöhung der Textmenge. Diese erhöht die extrinsische Belastung [8, 19]. Der „Expertise Reversal Effect“ sollte beim Video ebenfalls eine wesentliche Rolle spielen. Die Videos sollten zielgruppenspezifisch adressiert werden, um unnötige Informationen zu vermeiden [24]. Um DR1 zu adressieren ziehen wir das Signalisierungsprinzip heran. Informationen, die innerhalb des Videos besonders relevant sind, müssen hervorgehoben werden, um die Aufmerksamkeit des Nutzers zu wecken. Dadurch wird auch die Verarbeitung von weniger relevanten Informationen vermieden. Bei der Erstellung sollte ebenfalls das Redundanzprinzip berücksichtigt werden. Exakt die gleichen Informationen, die von den auditiven und visuellen Zentren empfangen werden, führen zu einer doppelten Beanspruchung. Informationen sollen sich ergänzen und nicht identisch sein. Des Weiteren sollte darauf geachtet werden, dass Abbildungen, zusammen mit der dazugehörigen Beschriftung präsentiert werden und somit das Prinzip der räumlichen Kontiguität eingehalten wird. Das Prinzip der temporalen Kontiguität sollte ebenfalls Beachtung finden. Zusammenhängende

Informationen sollten parallel bzw. nacheinander kommuniziert werden, um die kognitive Belastung bei der Verarbeitung zu reduzieren [8, 19, 22].

Awareness of Problem		Suggestion & Development	
Problem	**Design Anforderungen**	**Design Prinzipien**	**Gestaltungsrichtlinien an audiovisuelle Kommunikation**
Große Textmengen	DR1. Reduzierung der extrinsischen Belastung	1. Kohärenz Prinzip: Nicht relevante Information reduzieren 2. Signalisierung Prinzip: Relevante Informationen hervorheben 3. Redundanz Prinzip: Text bei Audioanimationen vermeiden 4. Räumliche Kontiguität Prinzip: Textzeilen neben der zugehörigen Grafik platzieren 5. Temporale Kontiguität Prinzip: Präsentation von visuellen und den dazugehörigen auditiven Inhalten parallel vorstellen	1. Gestalten Sie das Video möglichst simpel. Bei der Erstellung sollte der Fokus auf das Wesentliche gelenkt werden. 2. Wichtige Sachverhalte sollen klar identifizierbar und erkennbar sein. Hierfür ist es möglich Signalfarben zu wählen. Auf diese Weise wird die Aufmerksamkeit in die gewünschte Richtung gelenkt und das Arbeitsgedächtnis nicht unnötig belastet. 3. Vermeiden Sie die doppelte Darstellung identischer Informationen. Diese lenken ab und haben einen negativen Effekt auf das Arbeitsgedächtnis. Sich ergänzende Informationen oder Hervorhebungen sind jedoch förderlich für den Wissenstransfer. 4. Stellen Sie Inhalte übersichtlich im Video dar. Platzieren Sie hierfür die Beschriftung neben der zugehörigen Grafik. Auf diese Weise setzten sie die Kapazitäten des Arbeitsgedächtnisses beim Nutzer optimal ein. Die Informationen ergänzen sich. 5. Zeigen Sie zusammenhängende Inhalte parallel in Videos. Audiokommentare und die dazugehörenden Animationen müssen parallel im Video erscheinen, um im Arbeitsgedächtnis auf möglichst simple Weise zusammengefügt zu werden.
Schwierig zu verstehende Inhalte	DR2. Reduzierung der intrinsischen Belastung	6. Segmentierungsprinzip: Informationen in kleinere Segmente aufteilen	6. Teilen Sie die Informationen in kleinere Segmente auf. Komplizierte Sachverhalte, die aus diversen abhängigen Elementen bestehen, lassen sich so simpler kommunizieren. Ebenfalls wird das Arbeitsgedächtnis nicht überlastet.
Schwierig zu vermittelnde Inhalte	DR3. Erhöhung der lernbezogenen Belastung*	7. Personalisierungsprinzip: Informelle Informationsvermittlung wirkt sich positiver auf den Wissenstransfer aus als formelle	7. Kommunizieren Sie im Gesprächsstil. Durch die Nutzung einer einfachen informellen Sprache, hat die Zielgruppe das Gefühl an einem Gespräch teilzuhaben. Dies ist für die Motivation der Nutzer förderlich. Aus diesem Grund sollte die im Video verwendete Sprache nicht informell und distanziert sein. Dadurch lassen sich komplizierte Sachverhalte simpler darstellen. **insb. durch Schemata, da hierdurch das Arbeitsgedächtnis entlastet wird: Es müssen keine neuen Systematiken oder Raster gelernt werden.*

Tabelle 1.Design Prinzipien für die Reduzierung der kognitiven Belastung für die Steigerung der Kommunikation. Eigene Darstellung anhand [8, 19, 24, 25]

Als zweiten Punkt wird das Problem der „Schwierig zu verstehende Inhalte" adressiert. Dabei handelt es sich um eine intrinsische Belastung (DR2). Um dem Problem effektiv vorzubeugen, sollte das Segmentierungsprinzip herangezogen werden. Es beschreibt die Aufteilung von Informationen in kleinere Segmente. Das Arbeitsgedächtnis ist begrenzt, werden zu viele komplexe Informationen über einen längeren Zeitraum vermittelt, geht viel Informationsgehalt verloren. Daher ist es ratsam, mehrere Videos zu komplexen Sachverhalten zu erstellen[19, 22].

Als letztes Problem wurden „schwierig zu vermittelnden Inhalte" von den befragten Nutzern identifiziert. Das Problem lässt sich mit Hilfe der lernbezogenen kognitiven Belastung erklären (DR3). Diese sollte nach Möglichkeit gefördert werden, indem das Personalisierungsprinzip angewandt wird. Durch die Nutzung einer einfachen informellen Sprache, hat die Zielgruppe das Gefühl an einem Gespräch teilzuhaben. Dies ist für die Motivation der Nutzer förderlich. Aus diesem Grund sollte die im Video verwendete Sprache nicht informell und distanziert sein. Abgesehen davon lassen sich dadurch komplizierte Sachverhalte [8, 19, 20].

Zusammenfassend sollte festgehalten werden, dass die kognitive Belastung reduziert werden sollte und die lernbezogene Belastung zu fördern ist. Dies wird erreicht, indem die aufgestellten Design Prinzipien befolgt werden und der Lernende neue Schemata im Arbeitsgedächtnis aufbaut oder bereits vorhandene Schemata nutzt. Dies kann durch eine bessere Gestaltung des Lernmaterials geschehen. Die Benutzung und Förderung von Schemata wird als optimaler Weg beim Verständnislernen angesehen [22]. Die Gestaltung von interaktiven Videos umfasst neben kognitiven Aspekten ebenfalls visuelle Punkte. Bei der Erstellung von den Videos bildet das Raster eine gute Vorgehensweise, um diese zu strukturieren und zu standardisieren. Durch die Anordnung der Nutzungsfläche entsprechend der Rastervorgaben (Gestaltungslayout) wird eine Ordnung geschaffen, die dem Nutzer hilft vorhandene Schemata zu aktivieren[8, 24, 25]. Die Gestaltungsrichtlinien an audiovisuelle Kommunikation aus Tabelle 1 fassen nochmals die instanziierten Prinzipien zusammen. Somit wurden drei Mengen an problemspezifischen Designprinzipen identifiziert, die alle im Einzelnen auf jeweilige Richtlinien gemapped sind (siehe Nummerierung).

6 Diskussion und Ausblick

Die vorliegende Arbeit trägt zur Verbesserung der Kommunikation auf internen Crowdsourcing-Plattformen bei. In der bestehenden Literatur wurden Crowdsourcing und die Verbindung zur Unternehmensführung [26], unterschiedliche Arten von Crowdsourcing-Communities [27] oder Potentiale von Crowdsourcing [28] bereits betrachtet, nicht jedoch für Microlearning-Aufgaben. Es ist auffällig, dass verbal übermittelte Informationen innerhalb der Cognitive Load Theory eine besondere Bedeutung besitzen [24]. Erste Ansätze, die sich mit dem Thema eines effektiven Wissenstransfers auseinandersetzen, findet man in den von Mayer [2008] aufgestellten für die Wissensweitergabe aus kognitiver Sicht stärkenden Prinzipien, die für das Microlearning CSS herangezogen wurden [19]. Die Theorie der reduzierten sozialen Hinweisreize (Reduced-Social-Cues-Approach) vertritt die These, dass bei einer Reduzierung der Kommunikationskanäle ein Informationsverlust erfolgt, sodass bei reinen Textnachrichten im Unterschied zur sog. Face-to-Face-Kommunikation ein Teil der Sinneskanäle nicht genutzt werden. Dieser Umstand wird durch die Einführung der audiovisuellen Kommunikationsmöglichkeit erheblich optimiert. Die Kommunikation wurde als essentieller Bestandteil auf internen Crowdsourcing-Plattformen für die Erstellung von Microlearning-Lösungen identifiziert. Sie dient der Lösung von Problemen innerhalb eines Unternehmens, die durch Microlearninginhalte gelöst werden. Jedoch kann dies nur durch einen aktiven Austausch im CSS stattfinden. Die vorliegende Arbeit knüpft an den identifizierten Kommunikationsschwächen an. Diese wurden anhand von semistrukturierten Befragungen und mithilfe von CSS-Prozessen auf einer internen Crowdsourcing-Plattform erfasst. Anschließend erfolgte die Einbettung der Probleme in bestehende Theorien. Hieraus ergab sich eine Ableitung möglicher Herangehensweisen, die den Nutzer bei der Erstellung von Videos unterstützen.

Der theoretische Beitrag beschränkt sich auf ein noch nicht demonstriertes und evaluiertes, aber hierfür theoriegeleitetes Konzept zur Verbesserung des Crowdsourcing-Prozesses durch die Erstellung von Microlerninhalten. Dies geschieht in Form von Design Anforderungen und Design Prinzipien. Um die aus der Praxis identifizierten Probleme zu verstehen und einen ersten Schritt in eine Klasse von ungelösten Problemen [29] zu abstrahieren, wurden drei Design Anforderungen formuliert. Zum Adressieren dieser, wurden drei Mengen an zusammengehörenden

Design Prinzipien aus den 7 identifizierten Prinzipien nach Mayer [2008] entwickelt. Diese sollten übertragbar auf ähnliche Problemklassen (DR) sein. Der praktische Beitrag besteht darin, dass Plattformentwicklern, die sich darauf spezialisieren Probleme mit Microlerninhalten zu lösen, Prinzipien zur Gestaltung von Werkzeugen, die es ermöglichen diese konsumfreundlicher zu gestalten, zur Verfügung gestellt werden. Die konkreten Richtlinien sollen sowohl den Betreibern, als auch den Erstellern von Microlerninhalten auf Crowdsourcing-Systemen als Referenz dienen.

Um die erarbeiteten Prinzipien zu demonstrieren, muss im nächsten Schritt eine Anpassung der beobachteten Plattform an die Prinzipien geplant werden. Zur Evaluation kann ein Expertenworkshop mit Mitarbeitern des CSS-Anbieters, bei dem die erstellte Herangehensweise an die Nutzer weitergegeben wird, dienen. Die Entwicklung von Design Anforderungen und Design Prinzipien stellt zukünftiges Bestreben nach einer Designtheorie dar. Das Forschungsvorhaben der Autoren ist daher diesem zukünftigen Bestreben durch die Ergründung von Problemtheorien und Entwicklung des Problemraums zuzuordnen [18].

Literaturverzeichnis

1. Sultan, N. (2013): Knowledge management in the age of cloud computing and Web 2.0: Experiencing the power of disruptive innovations, in: International Journal of Information Management, Vol. 33, S. 160-165.
2. Li, M. M.; Peters, C. & Leimeister, J. M. (2016): Digitale Service-Systeme. In: Digitale Transformation im Unternehmen gestalten. Hrsg.: Gassmann, O. & Sutter, P. Verlag/Publisher: Carl Hanser Verlag, S. 29-38.
3. Bharati, P., Zhang, W. and Chaudhury, A. (2015): Better Knowledge with Social Media? Exploring the Roles of Social Capital and Organizational Knowledge Management, in: Journal of Knowledge Management, Vol. 19, Nr. 3, S. 456–475.
4. Döring, N. in Boos, M., Jonas, K. J. and Sassenberg, K. (Hrsg.): (2000). Computervermittelte Kommunikation in Organisationen. Göttingen: Hogrefe.
5. Vaishnavi, V. and Kuechler, B. (2004): Design Science Research in Formation Systems, http://www.desrist.org/design-research-in-information-systems/ Abruf: 10.10.2016.
6. Meth, H., Müller, B. and Maedche, A. (2015): Designing a Requirement Mining System, Journal of the Association for Information Systems, Vol. 16, No. 9, S. 799-837.
7. Hevner, A.R.; March, S. T.; Park, J.; Ram, S. (2004): Design Science in Information Systems Research, MIS Quarterly, Vol. 28, No. 1, S. 75-105.
8. Sweller, J., Ayres, P., Kalyuga, S. (2011). Cognitive Load Theory. New York, Dordrecht, Heidelberg, London: Springer Science + Business Media.
9. Howe, J. (2010). Crowdsourcing. Why the Power of the Crowd is Driving the Future of Business.
10. Leimeister, J. M. and Zogaj, S. (2013): Neue Arbeitsorganisation durch Crowdsourcing. EineLiteraturstudie, Hans-Böckler-Stiftung, Düsseldorf.
11. Zuchowski, O.; Posegga, O.; Schlagwein, D.; Fischbach, K. (2016): Internal Crowdsourcing: Conceptual Framework, Structured Review and Research Agenda, in: Journal of Information Technology (forthcoming).

12. Theo Hug, 2005, Micro Learning and Narration, fourth Media in Transition conference, May 6-8, 2005, MIT, Cambridge (MA), USA, http://web.mit.edu/comm-forum/mit4/papers/hug.pdf Abruf 08.10.2016.

13. Linder, M. (2006): Use These Tools, Your Mind Will Follow. Learning in Immersive Micromedia&Microknowledge Environments (Research Paper for ALT-C 2006: The Next Generation. Edinburgh, Scotland, Sept. 4–6, 2006).

14. Baumgartner, Peter. 2013. Warum gewinnt Microlearning zukünftig an Bedeutung? In: MicroLearning: Managing Innovation from Universities into Markets, hg. von Research Studios Austria Forschungsgesellschaft, 24–25. Salzburg.

15. Li, M. M. & Peters, C. (2016): Mastering Shakedown Through The User: The Need for User-Generated Services In Techno Change. In: European Conference on Information Systems (ECIS), Istanbul, Turkey.

16. Leimeister, J. M. (2012): Crowdsourcing. In: Zeitschrift für Controlling und Management (ZFCM), Jg. 56, Heft 6, S. 388-392.

17. Geiger, D.; Seedorf, S.; Schulze, T.; Schader, M.: Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes. Conference Paper, (2011)..

18. Venable, John R. "The role of theory and theorising in design science research." Proceedings of the 1st International Conference on Design Science in Information Systems and Technology (DESRIST 2006). 2006.

19. Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. American Psychologist , S. 760-769.

20. Mayer, R. E. (2001). The Case for Social Agency in Computer-Based Teaching: Do Students Learn More Deeply When They Interact With Animated Pedagogical Agents. Cognition and Instruction, S. 177-213.

21. Paivio, A. (1986). Mental representations: A dual coding-approach. Oxford University Press .

22. Clark, R., Nguyen, F., Sweller, J. (2006). Efficiency in learning. Evidence-based guidelines to manage cognitive load. San Francisco: CA: Pfeiffer.

23. Tinsdall-Ford, S. C. (1997). When Two Sensory Modes are Better Than One. Journal of Experimental Psychology , S. 257-287.

24. Sweller, J. (2003). Evolution of human cognitive architecture. The Psychology of Learning and Motivation , S. 215-266.

25. Wäger, M. (2010). Grafik und Gestaltung : das umfassende Handbuch. Bonn: Galileo Press.

26. Feller,J.; Finnegan, P.; Hayes, J.; O'Reilly, P. (2009): Institutionalising information asymmetry: governance structures for open innovation, in: Information Technology & People Vol. 22, Nr. 4, S. 297-316.

27. Kozinets, R. V., Hemetsberger, A. and Schau, H. J. (2008): The Wisdom of Consumer Crowds Collective Innovation in the Age of Networked Marketing, in: Journal of Macromarketing, Vol. 28, Nr. 4, S. 339-354.

28. Kleemann, F., Voß, G. G. and Rieder, K. (2008): Crowdsourcing und der Arbeitende Konsument. in: Arbeits- und Industriesoziologische Studien, Heft 1, S. 29-44.

29. Walls, J., Widmeyer, G.R. and El Sawy, O.A. (1992). Building an Information System Design Theory for Vigilant EIS, Information Systems Research, Vol. 3, No. 1, pp. 36-59.