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.

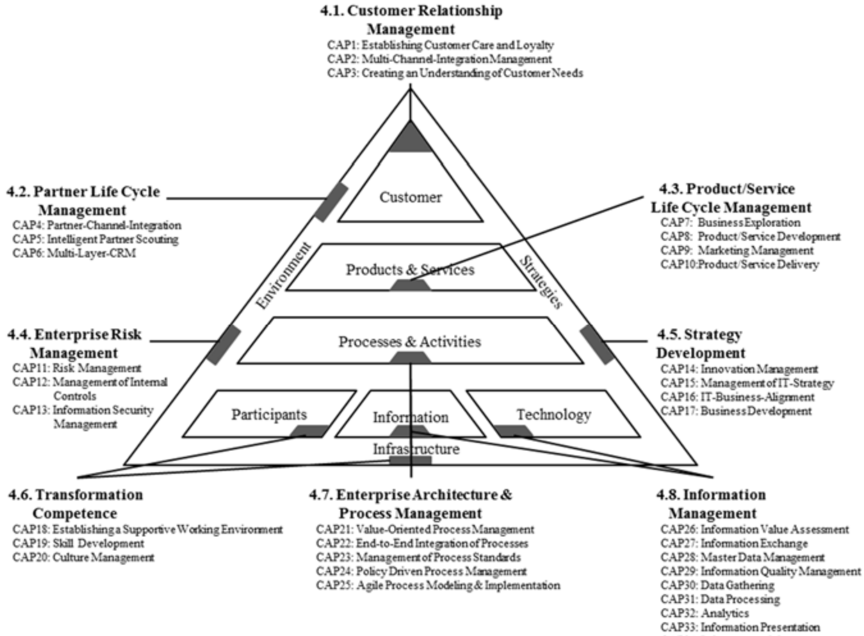

Figure 1. Big data analytics capability model according to the work system elements [42]

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).