AMCIS 2020 Proceedings

Data Science and Analytics for Decision Support (SIGDSA)

Aug 10th, 12:00 AM

Linking Big Data and Business: Design Parameters of Data-Driven Organizations

Janine Adina Hagen
LMU Munich, hagen@bwl.lmu.de

Thomas Hess
Institute for Information Systems, thess@bwl.lmu.de

Linking Big Data and Business: Design Parameters of Data-Driven Organizations

Completed Research

Janine Adina Hagen
LMU Munich
hagen@bwl.lmu.de

Thomas Hess
LMU Munich
t.hess@bwl.lmu.de

Abstract

Big data analytics is accepted to be an important driver of business value. However, this value does not come without a cost. Becoming a data-driven organization (DDO) necessitates a substantial transformation along the components structure, actors, task, and technology. Moreover, as successfully generating value from big data requires the utilization of data insights in business, attention needs to be assigned to the different actors from the data and business side, and their interrelation and collaboration. By taking a socio-technical systems perspective and utilizing a multi-case research approach, we developed a taxonomy to structure insights about different design parameters of a DDO. Thus, we contribute to the information systems literature by proposing a holistic design framework for DDOs paying tribute to its high collaboration requirements, and offer a compendium for managers with pathways how to design a DDO.

Keywords

Data-driven organization, big data analytics, collaboration, socio-technical system, taxonomy.

Introduction

Successfully deriving value from big data is a key success factor in the digitized world (Côrte-Real et al. 2019). The data-driven organization (DDO), characterized by a high level of data-based business decisions and actions (Schüritz et al. 2017), has been identified as a new type of organization and is predicted to grow 30% annually (Hopkins et al. 2018). However, organizations struggle to become data-driven as Gartner estimates that "80% of analytics insights will not deliver business outcomes through 2022" (White 2019). One reason for this is that companies do not manage to align their organizational models to the new requirements (Günther et al. 2017). Design choices concerning *structure*, *actors*, *tasks*, and *technology* need to be made, because a DDO can be seen as a holistic socio-technical system (Dremel et al. 2019). Specifically, as the actual business usage of data insights is a fundamental enabler to extract value from data (Côrte-Real et al. 2019), the creation of a big data analytics (BDA) ecosystem requires to envision the interrelation of the various actors, and to foster their collaboration (Pappas et al. 2018). In a DDO, at the level of frontline, these actors are data experts (e.g. data scientist, and data engineers) and business actors.

Scholars have identified several resources that enable organizations to create the necessary BDA capabilities, i.e. tangible resources (e.g. technology), human skills (business and data expertise) and intangible resources, e.g. data-driven culture (Gupta and George 2016). However, literature remains scarce on how these capabilities can be enabled and developed (Mikalef et al. 2020). Hence, the question arises of what a fruitful social subsystem (structure and actors) and technical subsystem (tasks and technology) looks like that enables these capabilities and nurtures the crucial data and business collaboration.

The *structural* component of the DDO has been examined in the literature focusing especially on whether to anchor BDA experts in a central, decentral, or hybrid manner. For instance, Dremel et al. (2017) examine the evolution of BDA at AUDI AG (centrally provided analytics-as-a-service), and Schüritz et al. (2017) propose design options for analytics competency centers. However, those studies focus on the central structure, and do not consider design alternatives. Other scholars discuss advantages and disadvantages of central, decentral and hybrid arrangements (e.g. Griffin and Davenport 2014): Centralizing data experts increases efficiency and expertise in company-wide BDA solutions, whereas the decentralized implementation ensures business proximity. However, they do not reveal how the stakeholders interact in such arrangements. Specifically, these studies do not consider the interrelation and collaboration of business and data experts (*actors* dimension), which is of special importance as business utilization of big data is key. It also remains unclear how the relevant *tasks* of the actors within the DDO work together to generate value. Concerning *technology*, scholars have identified requirements of a BDA infrastructure (e.g. Demchenko et al. 2014). However, we could not identify studies that consider how BDA technology is leveraged by the different actors. To the best of our knowledge, there are no studies that offer a holistic compendium of design parameters that provides several pathways to design a DDO, specifically considering the high demand for collaboration. Against this backdrop we pose the following research question (RQ):

RQ: *What are the design parameters of data-driven organizations?*

To answer this question, we developed a taxonomy based on information systems (IS) literature, and empirical research that compiles parameters and different manifestations for the design of a DDO. Specifically, we reveal approaches to link data and business actors within a DDO. The developed design compendium takes a holistic view as it a) pays tribute to the socio-technical character of a DDO and b) acknowledges the high need for collaboration by comprehensively considering business and data actors.

Foundations

Conceptual lenses: Socio-technical systems theory

In order to gain a holistic perspective on DDOs, we follow Dremel et al. (2019) by comprehending a DDO as a socio-technical system, as challenges in the implementation of BDA can result from structural, skill-, task-, and technology-related aspects (Vidgen et al. 2017). Socio-technical systems theory (STS) views organizations as multivariate systems with two subsystems including four interlinked components (Bostrom and Heinen 1977; Lyytinen and Newman 2008): The *social subsystem* includes the components structure and actors. *Structure* refers to project organizations and institutional arrangements. Main properties are e.g. the level of centralization, allocation of rights, and means of control. *Actors* are the organizational members and stakeholders who deliver the work. They can be described by e.g. their personal properties, skills, and differences. The *technical subsystem* in turn comprises the components tasks and technology. *Tasks* describe the way the work is delivered and the respective objectives. Main properties are e.g. task complexity and specificity. *Technology* includes the infrastructure and tools that are implemented. It includes a functional dimension (e.g. adaptability) and a systemic dimension (e.g. performance, ease of use). In order to function efficiently, the four components of the system must be harmonized (Lyytinen and Newman 2008). STS is used as a "lexicon for describing generative mechanisms and outcomes associated with the IS change" (Lyytinen and Newman 2008, p.594). In our case, this change is the anchorage of BDA in the organization to become a DDO. For our study, STS serves as conceptual lenses and frame for the taxonomy and thus ensures a holistic view. In the following, we will briefly characterize BDA and then outline the specifics of the DDO along the four socio-technical components.

Data-Driven Organization

A data-driven organization (DDO) "heavily relies on data to make decisions and take actions" (Schüritz et al. 2017, p.294). The data at the core of the DDO today often is called "big data" which can be delimited from "normal data" by the four V's v*olume*, *velocity, variety*, and *veracity* (Abbasi et al. 2016). However, for the purpose of this paper the characteristics of the data are subordinary, as the focus is more on turning data into business value (Gupta and George 2016). Thus, our understanding also includes conventional data (e.g. transactional data). As BDA can be regarded as a progression of earlier concepts such as business intelligence (BI) (Loebbecke and Picot 2015), a delimitation is necessary to carve out the DDO specifics: BI embraces a "system that helps support the decision-making process" (Frolick and Smith 2015, p.176), e.g. data warehousing. BI focusses on reporting, dashboards and backward analysis. In contrast, core to the DDO are three elements: forward-looking analyses focussing on next steps, management of the whole BDA value chain from data into action, and data-driven culture (Anderson 2015).

DDO structure. Whereas BI has mostly been organized in centers of excellence (e.g. Foster et al. 2015) that delivered analytics to the business units (BU), organizational literature suggests three models to anchor BDA in the organization (Grossman and Siegel 2014). These models must ensure collaboration between the involved big data and business stakeholders (Sharma et al. 2014). First, all data experts can be placed in a single unit (*centralized solution*). Scholars argue that those units are the most effective when they "integrate the work of centralized analysts with business units" (Griffin and Davenport 2014, p.5). Schüritz et al. (2017) mention the role of the analyst who translates BDA results into business value, however it is unclear how this translation would look like. *Decentralized structures* place data scientists close to business in the BUs. Troilo et al. (2017) argue that dispersed data scientists are an enabler to equip BUs with analytical models and an analytical mindset, but do not address how this can be executed. *Hybrid structures* place a critical mass of data scientists in a central unit, and the remaining ones throughout the organization. Scholars recommend this kind of structure as it can "address the need for [...] innovations" (Troilo et al. 2017, p.636) but again do not reveal options how collaboration within this structure can be fostered. Dremel et al. (2019) develop a portfolio management tool based on STS that we see as a valuable initial contribution how collaboration can be promoted. However, this tool was developed in a centralized DDO setting and thus again has a limited view within the structural dimension.

DDO actors. IS-literature differentiates data scientists, business analysts, and business users (Phillips-Wren et al. 2015). While data scientists need to have strong analytical and mathematical skills, they are also supposed to have business and communication knowledge (Troilo et al. 2017). On the business side, the analyst profile is characterized by a deeper data understanding than that of other business users, who are functional domain experts (Phillips-Wren et al. 2015). Analysts are also responsible for data interpretation, when the data scientists lack the necessaary business understanding (Schüritz et al. 2017).

DDO tasks: BDA is the whole process that extracts value from big data (Galbraith 2014) and includes data collection, preparation, storage, analysis, and usage (Phillips-Wren et al. 2015). In the following, we differentiate data- and business-oriented tasks within the DDO. The above mentioned big data experts are required to carry out data-oriented tasks such as descriptive analysis (reporting on the past), predictive analysis (future predictions), and prescriptive analysis (models to specify optimal actions) (Grover et al. 2018). However, deriving value from big data is only possible when the insights are used in the BUs "to weave a narrative making sense of the world" (Sharma et al. 2014, p.435). To make use of the data and create business value, related business tasks comprise organization performance improvement, process optimization, product and service innovation, and customer experience enrichment (Grover et al. 2018).

DDO technology. DDOs need an infrastructure that is capable to handle the great amount of data and facilitate the related analysis, which includes data repositories (on-premises or cloud) with high capacity (Halaweh and Massry 2015). More specifically, a BDA infrastructure requires cluster services, Hadoop related services and tools, special analytics tools, SQL / NoSQL databases, and MPP (massively parallel processing) databases (Demchenko et al. 2014). Scholars tend to favor a cloud-based data lake as it removes integration barriers and enables advanced analytics (Stein and Morrison 2014).

In summary, the literature focusses on selected DDO components (especially the structural) but provides less explanation of a holistic view. The necessity for collaboration is emphasized, but not addressed in terms of concrete solutions. Furthermore, literature does not detail how the technical infrastructure can be embedded into the social system of the DDO. Hereafter, we describe the taxonomy development that aims at compiling DDO design parameters and manifestations, equally considering the social and technical subsystem. Moreover, we aim to address the specific challenge of linking big data and business.

Methodology

Taxonomy Development

We opted for a taxonomy as this approach is proven for organizing knowledge (Wand et al. 1995), and such organization is currently lacking in terms of holistic design options for DDOs. We followed the method proposed for IS by Nickerson et al. (2013). This method aims to iteratively develop a taxonomy that sufficiently describes the object (DDO) by a set of dimensions, each consisting of a set of characteristics. Following Nickerson, we defined "design parameters of data-driven organizations" as the meta-characteristic that reflects the taxonomy purpose. We then determined three objective ending-conditions to terminate the taxonomy development: a) no new dimensions or characteristics were added in the last iteration, b) no dimensions or characteristics were merged or split in the last iteration, and c) every dimension is unique and not repeated. In contrast to Nickerson but in accordance with prior research (e.g.

Fuchs et al. 2019), we did not follow the objective ending condition that characteristics need to be mutually exclusive, but allowed for combinations of characteristics within one dimension. This turned out to be necessary to properly reflect the case study insights. Moreover, we ensured that subjective ending-conditions were met (taxonomy is concise, robust, comprehensive, extendible, and explanatory). To tie up to the literature, the first iteration was conceptual-to-empirical. We searched literature on IS and (data-driven) organization design as well as STS. We therefore came up with the first set of dimensions (design parameters of the DDO) and characteristics (manifestations of the design parameters) of the taxonomy that we grouped along the STS components structure, actors, technology, tasks. We then added insights from specific BDA literature which yielded the first taxonomy T_1.

Case Study Research

As literature does not provide a holistic, multi-dimensional view on DDOs including design options for data and business collaboration, we decided to employ an empirical-to-conceptual iteration as second cycle. We chose an exploratory multiple-case study research design as this is well suited to examine real life problems in depth in their natural context (Yin 2014). The objective was to extend and refine the taxonomy T_1 and to confirm its practical relevance. The acquired case sample includes five firms headquartered in Germany with heterogeneous business models and organizational structures (see table 1 for an overview). Cases were anonymized to ensure confidentiality. We first interviewed senior managers that oversee the DDO. The objective ending criteria were not met after this iteration, as these interviews delivered relevant practical insights as well as modifications and extensions of the initial taxonomy T_1, especially in the collaboration aspect. We therefore conducted a third iteration, where we discussed the insights from the second iteration (T_2) with employees from middle management, both data and business. As they did not claim major changes and only added marginal refinements to the taxonomy, we decided the objective ending-criteria to be met, resulting in the final taxonomy T_3. The number of interviewees ranged from three to six, depending on the organizational complexity. The numbers sorting the interviewees in table 1 will be used in the results section as quotation source. Data collection took place from 11/2019 to 01/2020. We conducted individual interviews that where primarily executed face-to-face and lasted from 30 to 75 minutes. All interviews were recorded, manually transcribed verbatim, and analyzed with ATLAS.ti. Data was triangulated with secondary data (e.g., firm websites, management reports) and internal data such as role and process descriptions. We performed two cycles of coding: The first cycle (descriptive coding) aimed at validating and refining T_1, and therefore included codes deducted from literature and STS (e.g. “central structure”). In the second cycle we ascribed emerging codes inductively (e.g. “data vs. human tournament”), whereas STS helped to assign them to the dimensions and characteristics (Miles et al. 2014).

Case	**Revenue** (2018)	**Number of interviews and roles of employees**	
“ManCo”	> 80 bn.€	6	1) Head of analytics and AI, 2) analytics manager (mgr.), 3) data transformation mgr., 4) marketing mgr., 5) data mgr., 6) data engineer
“MediaCo”	> 2 bn. €	4	1) Managing director, 2) head of data science and analytics, 3) senior business consultant, 4) senior mgr. BI
“RetailCo”	> 2 bn. €	4	1) Area mgr. IT, 2) area mgr. BI, 3) department head, 4) procurement analytics mgr.
“ToolCo”	> 1 bn. €	3	1) Head of data science, 2) data scientist, 3) scrum master
“PublicCo”	≈ 8 bn. €	3	1) Strategic digitalization mgr, 2) head of data science, 3) project mgr.

Table 1. Overview of case organizations and conducted interviews

Results

Before presenting and discussing the taxonomy, we briefly outline the cornerstones of the social and technical subsystems of the respective DDO to allow for a better grasp of the overall situation.

Case Descriptions

ManCo is an international manufacturer that is currently undergoing a wide-ranging data transformation program. Former structures have been grown without overarching guidelines, resulting in a decentralized structure with different standards across BUs. The target organization is a hybrid with a central data unit that steers the transformation across departments and has dedicated decentral contact persons to implement data management and governance in the BUs. In addition, there are data stewards assigned to BUs who are cross-department data experts with end-to-end responsibility for the data assets of their BUs. Although the focus is on decentral execution, a data lake secures centralized data processing. *MediaCo* is a tech and media company with several hundred brands and a deliberate decentral mindset. An IT service provider (100% subsidiary) offers software solutions as well as data science know-how to the parent group, focusing on those units with little or no BDA capabilities. The team consists of six employees that reflect the whole continuum from data to business. At MediaCo, data experts are anchored both decentrally in the bigger units as well as in the subsidiary, whereas the subsidiary also acts as service center for those units with no BDA capabilities (hybrid structure). Data is predominantly stored decentrally, however there is an initiative uniting similar business models on a joint data hub. *RetailCo* is a multichannel department store. BDA activities are organized centrally as part of the BI team. This “advanced analytics team” consists of two data scientists and two engineers and has a dual leadership (IT area manager and functional area manager). The functional area manager interacts with the BUs and bundles their requirements, whereas there are dedicated roles on the business side to connect with the central data unit. RetailCo aims at a data lake infrastructure, but is still facing the challenge of harmonizing historically grown silo systems. *ToolCo* is an international tool manufacturer that started BDA about three years ago. Focus has been on the delivery of prototypes for the BUs to show the benefits of BDA. That is why the small team consisting of a data scientist, engineer, and a scrum master is organized strictly centrally, led by a head of data science. In future there will be an in-between layer installed as BDA business partner to bring the BDA team and the BUs closer together. Some data is stored in the BUs, some data centrally. This infrastructure is seen to be inadequate for BDA due to restricting SLAs and is supposed to be adapted in future. *PublicCo* is a municipal utilities company where BDA is organized decentral in the BUs. The examined BDA team of the energy unit is the most mature, and consists of data scientists and software developers that work closely with the business. Collaboration is based on jointly developed use cases, and data is owned and stored within the BUs. As management is increasingly aware of the meaning of BDA for the value chain, they aim to build a group-wide data management function including a centralized data governance, BI, and acquisition team. The decentral data and business tandems shall be extended.

Design parameters of data-driven organizations

Subsequent we present the final taxonomy T_3 as a synthesis of IS literature and empirical research that we achieved after one conceptual-to-empirical and two empirical-to-conceptual iterations.

Subsystem/ component		Design dimension	Characteristic						
Social	Structure	*Anchoring of data experts*	Central	Hybrid	Decentral				
		Reporting line	Technology	Dual	Business				
		Horizontal linkage	Simplified examples	Meeting routines	Joint processes	Voluntary networks	Training	Integrator	Events
		Collaboration initiative	Business team	Data and business team	Data team (business need-based)	Data team (data-based)			
		Collaboration mode	Prototyping	Structured backlog	Occasional use cases				
		Control mechanisms	Business impact	Back testing	Utilization	Data vs. human tournament	Transformation KPIs		
	Actors	*Roles*	Data-oriented	Hybrid	Business-oriented				
Technical	Tasks	*Data tasks*	Descriptive analysis	Predictive analysis	Prescriptive analysis	Tools	Infra-structure	Visualization	Dashboards
		Business tasks	Process improvement	Decision-making	New insights	Products / services	Standardization		
	Tech-nology	*Data repository*	United	Hybrid	Dispersed				

Table 2. Taxonomy of design parameters of data-driven organizations

The taxonomy contains ten dimensions that are clustered along the STS components. We identified six dimensions of “structure” that are relevant for DDOs. *Anchoring of data experts* and *reporting line* refer to the BDA side, and *horizontal linkage*, *collaboration initiative, collaboration mode, and control* *mechanisms* specifically address the collaboration aspect of data and business. Except for the horizontal linkage dimension, all collaboration dimensions where added to the taxonomy based on our research results. The second component "actors" reflects the *roles* discussed in literature. The *task* component of the technical subsystem (except tools and infrastructure) is also grounded in literature as well as the technology component *data repository* which turned out to be crucial for the collaboration.

1) *Anchoring of data experts*. ToolCo and RetailCo have a *central structure* (e.g. Schüritz 2017), however show different business integration depths. To date, ToolCo has no established, joint processes with all the BUs, in contrast to RetailCo. However, they do not consider themselves as a center of excellence: "We consider ourselves 'business data science', we want to solve a business problem" (ToolCo, 1). The structure proved to be appropriate for the ramp up and for the delivery of prototypes. However, as the roadmap includes a professionalization of data-driven business, an in-between layer that serves as BDA business partner to bring data and business together is planned. RetailCo defends against the structure-immanent threat of alienation from business by designated counterparts in the BUs. The *hybrid design* (Grossman and Siegel 2014) allows for different nuances as well and is more than just the sum of central and decentral placement, since data experts have different tasks depending on their anchorage. At ManCo, the centralized team is responsible for transformation, governance and KPI management, whereas the central data team at MediaCo focusses on functional data science know how. MediaCo has the only BDA unit in the sample that acts as an own legal entity and is also pitching against externals. On the business side, decentrally placed data-oriented employees (e.g. data stewards) are either installed in all BUs (as a result of the strategy, ManCo) or just in selected BUs (depending on the unit size, MediaCo). *Decentral* data and business tandems (e.g. Troilo et al. 2017) work well for the heterogeneous BUs of PublicCo, although management strives for more standardization and is thus installing an additional central data function.

2) *Reporting line*. We observed three types of reporting lines (Grossman and Siegel 2014): *Technical reporting lines* go to the CIO (MediaCo and ManCo), *dual reporting lines* are drawn to CIO and CFO (RetailCo), and *business reporting lines* lead to the CPO (ToolCo) or functional area manager (PublicCo). Those decisions were either done intentionally (e.g. because of sponsorship of the respective C-level manager) or randomly due to personal data affinity of the board member (ToolCo, 1). Thus, literature could not be validated in saying that finding the right reporting line is a critical issue for DDOs (Grossman and Siegel 2014). However, literature could be validated in the fact that having a strong top-management support is an important organizational enabler (Troilo et al. 2017): "There needs to be sponsorship across the whole board and commitment to release necessary resources" (ManCo, 2).

3) *Horizontal linkage*. Structural overlays facilitate horizontal, cross-unit collaboration (Brown 1999). Our research revealed three specifications of collaboration in DDOs: collaboration between a) data and business, b) data and data, and c) between business and business, whereas the first one is the most crucial. The main objectives are "the creation of realistic expectations from business towards BDA and a basic technology understanding" (MediaCo, 2). Our research showed that *simplified examples* are applied frequently to show how to work with data and to "build an understanding of what is possible" (MediaCo, 1). These examples are mostly drawn from internal use cases. As mechanisms for formalized problem solving (Sting and Loch, 2016) we observed cross functional *meeting routines* that secure continuous alignment on shared projects (RetailCo, ManCo, PublicCo) and *joint processes* for data-driven assignments (PublicCo). As a more informal overlay, *voluntary networks* (Galbraith 1994) turned out to be beneficial for all three collaboration specifications: Data scientist networks in hybrid (MediaCo) or decentral organizations (PublicCo) secure ongoing exchange on similar data science problems and solutions, whereas data and business networks foster conversation about concrete use cases and economic potentials. A BDA related business-only network was also proposed: "We do something for one department, the other department does not know about it. Business people can learn from each other. Why don't we bundle their thoughts?" (ToolCo, 1). In addition, data-related *training* for the BUs ranged from hands-on training (MediaCo) which focusses on basics such as supervised and unsupervised machine learning to comprehensive training academies (ManCo) including mentoring and official degrees in different fields. Moreover, the data science head of ToolCo and the functional head of BI (RetailCo) act as *integration managers* (Galbraith 1994) to translate between data and business, e.g. by coordinating requirements. Occasionally, *events* such as company-wide summits are executed to gain momentum. (MediaCo).

4) *Collaboration initiative*. We introduce the design dimension "collaboration initiative" to describe the source (=actor) of the first impetus for a data-driven collaboration. At ManCo and ToolCo the *business team*

is the initiator of the collaboration aiming to solve a concrete business need. "Business always approaches me" (ToolCo, 1). Depending on the use case and data ownership, collaboration initiative can also come mutually from the *business and data team*: "In the majority of cases the data owner [the BU] approaches us. We are the driving force when two things came together, the know-how and data access" (MediaCo, 1). RetailCo and PublicCo push collaboration mostly from the *data team*. "You have to beg a bit to get the appointment to make life easier for them" (RetailCo, 2). This push is mostly *based on business needs*, when there is sufficient knowledge of the business model within the data team, but in specific cases also *based on data*. However, impulses based on pure data patterns may lead to a dead end: "Often a scientist says 'we noticed something noticeable', where the business says 'yes, but it is in the nature of things'" (MediaCo, 3).

5) Collaboration mode. We define the design dimension "collaboration mode" as the method by which the data and business collaboration is organized. Most prevalent is the development of *prototypes* (ToolCo, MediaCo, PublicCo, RetailCo) that is supposed "to help business to evaluate the potential of an operative implementation" (MediaCo, 1). We found that the more the data-business collaboration is established, the more *structured backlogs* are used (e.g. RetailCo, PublicCo): "The first step is to build a roadmap and have a lot of problems that you evaluate commercially" (PublicCo, 2). Whereas PublicCo's backlog only contains BU-internal use cases, the central data team of ManCo has collected use cases form all BUs and prioritized them regarding their economic impacts for the company and allocates resources accordingly. *Occasional use cases* are another collaboration mode representing the sporadic collaboration based on different application scenarios and objectives (ToolCo, MediaCo). Eventually it can be concluded that collaboration between data and business in DDOs is founded on "giving and taking" (MediaCo, 2).

6) Control mechanisms. The sample revealed five instruments to control the success of the DDO (Ouchi 1979), which in all cases included the data and the business side: Measuring the *business impact*, the effect of BDA techniques on defined business KPIs, is the most frequent control mechanism used by all case organizations. The measured impact can be top and bottom line, but also workplace-related like simplification and acceleration of tasks. *Back testing* means the assessment of how a data model would have performed ex-post, e.g. the retrospective prediction of revenues based on a point in history. The actual *utilization* of data tools provided to the business is another way how DDO success is measured. "Fairly easy to evaluate, if the manager used the application, or not" (ToolCo, 1). Another control mechanism that is also applied to convince BUs is the *data vs. human tournament*: "If we substitute things, we continue the manual work for half a year, and our AI is conducted in parallel and we compete to see who is better, and then business decides whether they want to use it" (MediaCo, 2). In the case of ManCo, DDO *transformation KPIs* are also applied; for example, the amount of data assets generated.

7) *Roles*. Our research reveals three role clusters (Phillips-Wren et al. 2015) that cover the whole BDA value chain: On the *data-oriented* end of the continuum with less business focus, the most prevalent roles (all cases) are data scientists and data engineers, who often work in tandem together (ToolCo, MediaCo, RetailCo). "I compare it with a country conquest. Data scientists are like pioneers, they are the first to go to the country and see what it looks like. Data engineers are the ones who build houses" (RetailCo, 2). Other roles in this cluster are data architects and frontend developers (RetailCo) as well as software developers (PublicCo). Some data scientists at PublicCo do not have a dedicated education in this field, but migrated into this profession from business, thus having both data and business competencies. All interviewees share the opinion that the absence of business know how of centralized data experts yields inferior results as "even data heroes will never reach excellence if they don't understand the business" (PublicCo, 2). Having all three skills (science, engineering, business) "would be the ideal situation which is scarce at the job market" (MediaCo, 2). An approximation to this profile are *hybrid roles* (Schüritz et al. 2017) that are prevalent in all central and hybrid structures represented in the sample, albeit with different manifestations: We could observe business consultants (MediaCo), scrum masters (ToolCo), a functional BI manager (RetailCo), and data stewards (ManCo). These roles combine a basic understanding of data and technology but also business and communication and act as "translator" between the involved parties. "Without me, there would be a lot of misunderstanding. Effort would be put into things that do not deliver the requested result" (MediaCo, 3). Thus, this role shares some features of the horizontal coordination mechanisms of the integrator. PublicCo lacks this role as the head of data science sees himself as business: "I'm 100% part of it, I'm not the external one who is 'kind of weird' " (PublicCo, 2). Lastly, *business-oriented roles* comprise requesters and users of data solutions that are in most cases not formalized. RetailCo is the only case that makes a distinction at the business side between "power users" who can apply their own data tools, "key users" who can work with the big data interface and "users", who still prefer excel over BDA tools.

8) *Data tasks*. Two of the three types of BDA tasks (Grover et al. 2018) could be confirmed: *Descriptive* and *predictive analysis* are applied by all organizations. Examples include the prediction of incoming goods (RetailCo), and maintenance dates (PublicCo). The application of *prescriptive analysis* has not been reported to be actively applied, but should be mentioned for the sake of completeness as it has been proven to be increasingly important (Grover et al. 2018). Beyond these three tasks, our research showed that selected teams are also responsible for building *tools* for the BUs, e.g. to substitute former tools like excel (PublicCo, ManCo), for building relevant *infrastructure*, e.g. data warehouses ("we also do a lot of basic work, first of all merging data silos so that we can analyze the data" (MediaCo, 2), and to *visualize* data in new ways (e.g. with tableau). Another (originally BI-related) basis data task is the provision of *dashboards*, that "bring the data together and show business what is going on" (MediaCo, 2).

9) *Business tasks*. In line with the literature (Grover et al. 2018) we revealed *process improvement* as a task prevalent in all cases. This ranges from minor work process improvements ("we see that there is great pain in the BUs and lots of potential in the small things", ManCo, 2) to core processes (e.g. predictive maintenance at PublicCo). The second most widespread business task is the improvement of *decision-making* (e.g. Dremel et al. 2019) to "replace gut feelings with a forecast model" (RetailCo, 1). This is also facilitated by *new insights* that can be generated due to a combination of former isolated data sources. The development of *data-driven products and services* (e.g. Grover et al. 2018) up to complete new business models is acknowledged to be of high potential (MediaCo) and partially in preparation: "The last piece we want to look into is where we sell data products" (ToolCo, 1). At RetailCo, BDA is also utilized for *standardization* of former non-orchestrated silo solutions across departments ("we don't have to explain to the CEO anymore why units report different KPIs values for the same period because we all look at the same reports", RetailCo, 4). This standardization task can be most likely connected to the category of organization performance improvement (Grover et al. 2018).

10) *Data repository*. Core of the technical DDO infrastructure is a platform for "collecting, integrating, sharing, processing, storing, and managing big data" (Grover et al. 2018, p.399). We identified three manifestations that we label united, hybrid, and dispersed. The cases revealed that the implemented alternative does not necessarily correspond with the organizational structure: Crucial in ManCo's transformation strategy towards a hybrid structure is the implementation of a company-wide data lake (Stein and Morrison 2014) that we call a *united data repository*. It consists of three layers: Data is injected into the source layer and then prepared by data engineers on the prepare layer to reach a unified quality. On the semantic layer, data assets are built by a tandem of data engineers and stewards. "This way we make data easily available for all BUs" (ManCo, 2). In contrast, MediaCo has a *hybrid data repository* consisting of united and dispersed data storages: The data team receives data access mostly only for the assignment as data is predominantly stored in the brands. However, they built a central data hub where brands with similar business models can share data. In some cases, the data team also owns data assets itself. Centrally organized RetailCo is still harmonizing historically grown, decentral silo systems (*dispersed repository*) but also aims at a data lake. ToolCo's data repository is also dispersed, and they report issues getting access to relevant data from different sources which are restricted due to outdated policies: "Back then, nobody thought about big data" (ToolCo, 1). This makes the data team the "bottle neck" in the collaboration (ToolCo, 1). The data of PublicCo is also stored decentrally in the BUs who have the ownership and are responsible for data security. However, they did not report any challenges as data is very specific for every BU.

Conclusion, Limitations and Suggestions for Future Research

Referring to our research question we found ten parameters that are relevant when designing a DDO. These are 1) anchoring of data experts, 2) reporting line, 3) horizontal linkage, 4) collaboration initiative, 5) collaboration mode, 6) control mechanisms, 7) roles, 8) data tasks, 9) business tasks, and 10) data repository. As all dimensions have different manifestations, we can state that companies design the social and technical subsystem of their DDO in different ways. However, we see a slight tendency for a hybrid setting in the social subsystem (central unit for overall coordination of data-driven initiatives and decentral data experts to pursue business value realization together with the BUs), and a centralized setting in the technical subsystem (single, company-wide data repository for cross-unit insights). Notwithstanding the chosen design, all companies apply dedicated collaboration approaches to link big data and business actors.

First, our study contributes to academia by developing an initial, holistic taxonomy to reveal different possible design parameters for DDOs along all components of a socio-technical system, integrating both the business and data perspective and their interrelations. Thus, we add to IS literature that has primarily been focusing on selected organizational design aspects of the DDO and contribute to a better understanding of the different cornerstones that need to be tackled. In particular, we extend literature by revealing collaboration approaches for big data and business by bringing approaches for horizontal linkage, collaboration initiative and mode, and control mechanisms to light. These approaches are initial solutions to the identified research gaps concerning how to integrate the work of data experts with business units. Moreover, this work establishes a basis (from an organizational design perspective) to enable BDA capabilities. Second, the taxonomy serves as a compendium for managers, presenting possible pathways they can take. A major practical benefit of the taxonomy is the integrative view on both business and data stakeholders and their collaboration which prepares the ground for successful business utilization of data insights. The taxonomy is especially applicable in early design stages, either in a DDO transformation program, or as a starting point for a new DDO, securing that no important parameters will be disregarded and that the integrative view of business and data is maintained. However, the final design depends on the specific contingencies and experiences of the organization. The taxonomy is not intended to be a DDO-maturity checklist, as the sole prevalence of a parameter does not tell anything about the BDA performance.

Finally, the taxonomy is an initial attempt to holistically organize DDO design knowledge with special focus on business and data collaboration and thus a basis for further research: In order to be even more applicable for managers, relations between the dimensions and characteristics should be identified by qualitatively and quantitatively investigating more cases. The taxonomy thus can serve as a basis for the deduction of a DDO typology. In this context, it would also be fruitful to examine contingencies which organizations face and drivers of the respective DDO type, which would support managers in making the right design decisions. Lastly, as the literature provides scant insight on metrics for taxonomy usefulness, we wish to emphasize that this taxonomy may develop over time and should be further validated.

References

Abbasi, A., Sarker, S., and Chiang, R. H. 2016. "Big Data Research in Information Systems: Toward an Inclusive Research Agenda," *Journal of the Association for Information Systems* (17:2), pp. i - xxxii.

Anderson, C.A. 2015. *Creating a Data-Driven Organization: Practical Advice from the Trenches*, Sebastopol, USA: O'Reilly Media.

Bostrom, R. P., and Heinen, J. S. 1977. "MIS Problems and Failures: A Socio-Technical Perspective. Part I: The Causes," *MIS Quarterly* (1:3), pp. 17-32.

Brown, C. V. 1999. "Horizontal Mechanisms under Differing IS Organization Contexts," *MIS Quarterly* (23:3), pp. 421-454.

Côrte-Real, N., Ruivo, P., Oliveira, T., and Popovič, A. 2019. "Unlocking the Drivers of Big Data Analytics Value in Firms," *Journal of Business Research* (97), pp. 160-173.

Demchenko, Y., de Laat, C., and Membrey, P. 2014. "Defining Architecture Components of the Big Data Ecosystem", *Proceedings of 2014 International Conference on Collaboration Technologies and Systems (CTS),* Minneapolis, USA.

Dremel, C., Stoeckli, E., Wulf, J., and Brenner, W. 2019. "A Socio-Technical Approach to Manage Analytics-as-a-Service–Results of an Action Design Research Project," *Proceedings of the 14th International Conference on Wirtschaftsinformatik*, Siegen, Germany.

Dremel, C., Wulf, J., Herterich, M. M., Waizmann, J.-C., and Brenner, W. 2017. "How AUDI AG Established Big Data Analytics in Its Digital Transformation," *MIS Quarterly Executive* (16:2), pp. 81-100.

Foster, K., Smith, G., Ariyachandra, T., and Frolick, M. N. 2015. "Business Intelligence Competency Center: Improving Data and Decisions", *Information Systems Management* (32:3), pp. 229-233.

Frolick, M. N., and Smith, G. 2015. "From the Special Issue Editors: Business Intelligence", *Information Systems Management* (32:3), p. 176.

Fuchs, C., Barthel, P., Herberg, I., Berger, M., and Hess, T. 2019. "Characterizing Approaches to Digital Transformation: Development of a Taxonomy of Digital Units," *Proceedings of the 14th International Conference on Wirtschaftsinformatik*, Siegen, Germany.

Galbraith, J. R. 1994. *Competing with Flexible Lateral Organizations* (2nd ed.), Boston, USA: Addison-Wesley.

Galbraith, J. R. 2014. "Organizational Design Challenges Resulting from Big Data," *Journal of Organization Design* (3:1), pp. 2-13.

Griffin, J., and Davenport, T. 2014. "Organizing Analytics - from the inside Out," Deloitte, pp. 1-8.

Grossman, R., and Siegel, K. 2014. "Organizational Models for Big Data and Analytics," *Journal of Organization Design* (3:1), pp. 20-25.

Grover, V., Chiang, R. H., Liang, T.-P., and Zhang, D. 2018. "Creating Strategic Business Value from Big Data Analytics: A Research Framework," *Journal of Management Information Systems* (35:2), pp. 388-423.

Günther, W. A., Mehrizi, M. H. R., Huysman, M., and Feldberg, F. 2017. "Debating Big Data: A Literature Review on Realizing Value from Big Data," *The Journal of Strategic Information Systems* (26:3), pp. 191-209.

Gupta, M., and George, J. F. 2016. "Toward the Development of a Big Data Analytics Capability," *Information & Management* (53:8), pp. 1049-1064.

Halaweh, M., and Massry, A.E. 2015. "Conceptual Model for Successful Implementation of Big Data in Organizations," *Journal of International Technology and Information Management* (24:2), pp. 21-34.

Hopkins, B., McCormick, J., and Schadler, T. 2018. „Insights-Driven Businesses Set The Pace For Global Growth: The Vision Report Of The Insights-Driven Business Playbook, Retrieved 01/2020, from https://www.forrester.com/report/InsightsDriven+Businesses+Set+The+Pace+For+Global+Growth/-/E-RES130848.

Loebbecke, C., and Picot, A. 2015. "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), pp. 149-157.

Lyytinen, K., and Newman, M. 2008. "Explaining Information Systems Change: A Punctuated Socio-Technical Change Model," *European Journal of Information Systems* (17:6), pp. 589-613.

Mikalef, P., Pappas I.O., Krogstie, J., and Pavlou, P.A. 2020. "Big Data and Business Analytics: A Research Agenda for Realizing Business Value," *Information & Management* (57:1), 10323.

Miles, M. B., Huberman, A. M., and Saldana, J. 2014. *Qualitative Data Analysis: A Methods Sourcebook* (3rd ed.), Thousand Oaks: SAGE Publications.

Nickerson, R.C., Varshney, U., and Muntermann, U. 2013. "A Method for Taxonomy Development and its Application in Information Systems," *European Journal of Information Systems* (22:3), pp. 336-359.

Ouchi, W. G. 1979. "A Conceptual Framework for the Design of Organizational Control Mechanisms," *Management Science* (25:9), pp. 833-848.

Pappas, I.O., Mikalef, P., Giannakos, M.N., Krogstie, J., and Lekakos, G. 2018. "Big Data and Business Analytics Ecosystems: Paving the Way Towards Digital Transformation and Sustainable Societies," *Information Systems and e-Business Management* (16:3), pp. 479–491.

Phillips-Wren, G. E., Iyer, L. S., Kulkarni, U. R., and Ariyachandra, T. 2015. "Business Analytics in the Context of Big Data: A Roadmap for Research," *Communications of the Association for Information Systems* (37:23), pp. 448 – 472.

Schüritz, R., Brand, E., Satzger, G., and Bischhoffshausen, J. 2017. "How to Cultivate Analytics Capabilities Within an Organization?–Design and Types of Analytics Competency Centers," *Proceedings of the 25th European Conference on Information Systems*, Guimarães, Portugal.

Sharma, R., Mithas, S., and Kankanhalli, A. 2014. "Transforming Decision-Making Processes: A Research Agenda for Understanding the Impact of Business Analytics on Organisations," *European Journal of Information Systems* (23:4), pp. 433–441.

Stein, B., and Morrison, A. 2014. "The enterprise data lake: Better integration and deeper analytics", *PWC Technology Forecast: Rethinking integration* (1), pp. 1-9.

Sting, F. J., and Loch, C. H. 2016. "Implementing Operations Strategy: How Vertical and Horizontal Coordination Interact," *Production and Operations Management* (25:7), pp. 1177-1193.

Troilo, G., De Luca, L. M., and Guenzi, P. 2017. "Linking Data-Rich Environments with Service Innovation in Incumbent Firms: A Conceptual Framework and Research Propositions", *Journal of Product Innovation Management* (34:5), pp. 617-639.

Vidgen, R., Shaw, S., and Grant, D. B. 2017. "Management challenges in creating value from business analytics", *European Journal of Operational Research* (261:2), pp. 626-639.

Wand, Y., Monarchi, D. E., Parsons, J., and Woo, C. C. 1995. "Theoretical Foundations for Conceptual Modelling in Information Systems Development," *Decision Support Systems* (15:4), pp. 285-304.

White, A. 2019. "Our Top Data and Analytics Predicts for 2019." Retrieved 12/2019, from https://blogs.gartner.com/andrew_white/2019/01/03/our-top-data-and-analytics-predicts-for-2019/.

Yin, R. K. 2014. *Case Study Research: Design and Methods*, Thousand Oaks, USA: SAGE Publications.