Data as Strategic Resources

Gunther, W.A.

2019

CHAPTER 1

INTRODUCTION

Ask different people what “data” means to them and you will most probably get very different answers. To scholars, for example, data are recordings of empirical observations that have been rigorously collected, based on which they aim to draw solid conclusions. To computer scientists and engineers, data may represent digital records that need to be stored in such a way that they are accessible and can be used by various systems and applications. Those interested in artificial intelligence—or science fiction—may associate the term with machine learning and androids. But ask “managers” what data means to them, and they might suddenly start talking about how crucial data are to their business and how they will use data to achieve a competitive advantage. What is that about?

Over recent years, managers have been triggered to think about different ways in which data, and the analysis thereof, can help their business. Many press articles, business journals, industry magazines, and whitepapers have popularized terms such as “big data” and “data science” for organizations (Madsen & Stenheim, 2016). Managers are constantly being confronted with memorable examples of organizations that effectively exploit data: how Netflix is relying on data to make recommendations and inform the contents of new series and movies (Lycett, 2013; Varian, 2014); how Google is making predictions based on search queries (Mayer-Schönberger & Cukier, 2013; Choi & Varian, 2012), and how Facebook has been making money based on data-driven advertising (Marr, 2016). Managers who take note of these examples and trends may feel that they cannot fall behind and are triggered to explore the opportunities that data can offer to their organizations.

Consequently, organizations have been jumping on the data-bandwagon. Institutes such as Gartner and The International Data Corporation illustrate that organizations have increasingly been spending their resources on data and analytics, and predict that this spending will only grow in the coming years (Gartner, 2013; Gartner, 2018-1; International Data Corporation, 2018). We see examples of data-driven solutions emerging in healthcare (Wang & Hajli, 2017), banking (Bholat, 2015), and education (Cech et al., 2015). We see that organizations are infusing data into their existing ways of creating and appropriating value (Schüritz & Satzger, 2016), and are creating whole new, data-driven business models (Hartmann et al., 2016; Loebbecke & Picot, 2015). This all signifies a trend wherein organizations are starting to see data not only as inputs to their enterprise systems or as by-products of their IT activities, but as strategic resources that they can leverage to “create differential value” (Bharadwaj et al., 2013, p. 472).

All of the optimism and hopes around data as strategic resources are, however, also accompanied by skepticism. For example, scholars have argued that the potential of data as strategic resources may be overshadowed by hypes, buzzwords, and “hullabaloo” (e.g., Arnott & Pervan, 2014; Boellstorff, 2013; Madsen & Stenheim, 2016). Moreover, anecdotal evidence shows that managers struggle to effectively implement and utilize data and related technologies (Gartner, 2016; Ransbotham et al., 2016), and "are slow to advance in data and analytics" (Gartner, 2018-2). This suggests that we are still a long way from understanding how data may actually be leveraged as strategic resources.

In this dissertation, I aim to move beyond the hype that is driven by vendors and management gurus, and critically assess how organizations approach this phenomenon. I perform three studies in which I draw on various methods in order to understand: How do organizations explore the opportunities of data as strategic resources? Before introducing the different studies, I will first conceptualize "data", consider different ways in which data may be leveraged strategically according to the literature, and reflect on what value organizations hope to realize by doing so.

1.1 Conceptualizing Data

I began my introduction by arguing that the term "data" can mean different things to different people. Numerous definitions of data can also be found in the literature[1]. When I refer to data in this dissertation, I generally refer to records that have been created (by humans or machines) at specific points in time and are stored somewhere (e.g., in databases) in digital forms. It is mainly these dematerialized, digital records that recent discourse has been promoting and that organizations nowadays aim to leverage as strategic resources (Lycett, 2013).

As digital records, data are "editable", "interactive", "reprogrammable", and "distributable" (see Kallinikos et al., 2013). Data are "editable" as they may be changed over time when organizational actors work with the data in practice (Ekbia, 2009; Kallinikos et al., 2013). For example, employees and customers can—through the use of database queries and other digital artifacts—extend, delete, and update (parts of) the data. Actors may also interact with the data without editing them, for example, by zooming in on certain parts of the data and working with representations of the data (Leonardi, 2010). In doing so, different actors may assign different meanings to the data (Jones, 2018). Data are "reprogrammable" as their logical structure can be changed, and are "distributed" as they do not typically reside in one place (Kallinikos et al., 2013). In other words, data are continuously shaped, interpreted, and used by distributed actors within their social contexts. Recognizing this, I take a socio-technical perspective and consider both technical and non-technical characteristics of data when studying how organizations explore their strategic potential.

Nowadays, organizations often talk about "big data", a term that gained popularity some years into the new century (Chen et al., 2012; Arnott & Pervan, 2014). While this term can also mean different things to different peo-

1 See, for example, Zins (2007) for an overview of 45 definitions.

ple (Kitchin & McArdle, 2016), it is often associated with large volumes of data coming from various sources and in many different forms, and that are generated, collected, and analyzed at increasing speeds (Laney, 2001). Importantly, with the rise of (the term) big data also arose the conviction that data could potentially be strategic resources (Davenport et al., 2012; Madsen & Stenheim, 2016). In the following sections, I will explain how data may be leveraged as strategic resources and what value organizations hope to realize by doing so.

1.2 Leveraging Data as Strategic Resources

To leverage data as strategic resources means to use data in ways that might eventually result in some kind of strategic "value" (Bharadwaj et al., 2013). In this dissertation, I am open to different ways in which data can be leveraged as such. Two common ways are, for example, relying on data to inform decision making and leveraging data in the creation of data-driven products and services (e.g., Dallemule & Davenport, 2017; Davenport et al., 2012; Wixom & Ross, 2017; Woerner & Wixom, 2015).

1.2.1 Data-Informed Decision Making

One way in which organizations can leverage data strategically is by relying on data to inform decision making. Both scholars and practitioners seem to believe that the adoption of (big) data and data-related technologies can help organizational actors at all levels to gain insights that lead to "better" decisions. For example, papers have suggested that those who rely on data and the rigorous analyses thereof consistently make better decisions than those who do not (Brynjolfsson et al., 2011; Davenport, 2006; Kiron et al., 2014). Similarly, literature has argued that organizations that are able to establish a culture of data-driven decision making are able to run their business more "intelligently" than those that do not (Davis, 2014; LaValle et al., 2012). One of the most influential works on this topic even argues that: "Data-driven decisions are better decisions—it's as simple as that" (McAfee & Brynjolfsson 2012, p. 63). Although a stream of literature is emerging that contests this assumption (e.g., Sharma et al., 2014; Shollo & Galliers, 2015; Pachidi et al., 2014), statements like these signal a conviction that data can be of strategic value when they are used for decision making.

Organizations may rely on data to inform both operational and strategic decisions (Holsapple et al., 2014; Sharma et al., 2014; Lycett, 2013). At the level of operations, for example, employees may use data about the organization's supply chains to evaluate their performance and better manage inventories and market demands (Chae et al., 2014; Chen et al., 2015; Davenport, 2006). Employees may also use data to interact with customers, continuously update their pricing methods, and analyze and minimize operational risks (Dallemulle & Davenport, 2017; Davenport, 2006; McAfee & Brynjolfsson, 2012; Wixom & Ross, 2017). Organizations may even automate operational decision making by relying on data in combination with algorithms, in order to operate more efficiently and effectively (Markus, 2015; Newell & Marabelli, 2015). At the strategic level, senior managers may use data to, for example, help them inform their business strategies; guide them in their merger and acquisition processes; forecast trends and identify disruptors, and generally look for new business opportunities (Constantiou & Kallinikos, 2015; Kiron et al., 2014; Lau et al., 2012; Van Rijmenam et al., 2018).

1.2.2 Data-Driven Products and Services

Another way in which organizations can leverage data strategically is in the creation of products and services (Davenport et al., 2012; Davenport & Kudyba, 2016). For example, organizations may "wrap" data around existing offerings, by taking products and services that are not initially data-driven and finding ways to use data to increase their value (Woerner & Wixom, 2015; Wixom & Ross, 2017). Quite a number of examples of such wrapping can be found in practice. Walt Disney, for example, has introduced wristbands that can track visitors' movements and store data about them, in order to give customers a more "magical" experience—e.g., Disney characters can use the gathered information to personally greet children (Marr, 2016). In fact, under the heading of the "Internet of Things" (Madakam et al., 2015), organizations are increasingly equipping their products with sensors that can collect all kinds of data about their usage. These data can be analyzed (in real-time) and give consumers insights to enrich their product experiences (Woerner and Wixom, 2015; Wixom & Ross, 2017). For example, ski equipment manufacturer Rossignol has, in collaboration with a startup, recently released a type of sensor that can collect and analyze a person's skiing style in real-time, to help them improve their skills (Rossignol, 2019).

Organizations may also treat the data themselves as core products (Hartmann et al., 2016; Woerner & Wixom, 2015; Wixom & Ross, 2017). When the data that organizations have (or can gain) access to are of high quality, they may try to monetize these data by selling them as standalone products (Woerner & Wixom, 2015; Wixom & Ross, 2017). Organizations that do this are typically called data providers, vendors, or "brokers" (Brooks, 2005). On top of selling the "raw" data, organizations may also sell the insights that they gain from the data (Hartmann et al., 2016). For example, organizations that have all kinds of data about consumers may analyze these data to help other organizations find their target groups.

1.3 From Data-driven Decisions, Products, and Services to Value

The key assumption is that, when organizations make data-informed decisions and create data-driven products and services, this will eventually result in "value" (Kiron et al., 2014; Seddon et al., 2017; Sharma et al., 2014). What this value entails in practice, and how organizations measure it, may strongly depend on the type of organization, their strategy, and the specific goals for which they rely on data (see Grover et al., 2018; Tempini, 2017; and Wamba et al., 2015 for different types of value). At the most basic level, we may consider "economic value" and "social value" (Emerson, 2003).

Economic value can be measured by an organization's increase in profit, business growth, and competitive advantage resulting from the adoption of big data and data-related technologies. Generally, organizations that rely on data and "big data" to guide their strategies and operations are expected to perform better than organizations that do not (LaValle et al., 2011; McAfee & Brynjolfsson, 2012). For example, McAfee & Brynjolfsson (2012) found, based on interviews with executives and an examination of the organizations' performance data, that "[t]he more companies characterize themselves as data-driven, the better they performed on objective measures of financial and operational results" (p. 64). This conclusion seems to be generally accepted; when organizations can efficiently and effectively compete on data and data analytics (Davenport, 2006), they will achieve greater profits (e.g., Tyagi, 2003).

Organizations may also (simultaneously) aim to create social value. Social value creation can be defined as "the creation of benefits or reductions of costs for society—through efforts to address social needs and problems—in ways that go beyond the private gains and general benefits of market activity" (Phills et al., 2008, p. 39). Although this definition puts emphasis on the benefits for society, organizations may also focus on preserving and enhancing the lives of individuals (Auserwald, 2009). Numerous examples have been provided in the literature that promote the strategic use of data for creating social value. In healthcare, for example, data and data-related technologies may be deployed to detect diseases and provide personalized care (e.g., Raghupathi & Raghupathi, 2014; Wang & Hajli, 2017). In education, schools may analyze data to gain insights that can help improve students' learning performance (e.g., Cech et al., 2015; Long & Siemens, 2011; Ferguson, 2012). Data may even be leveraged by intelligence and security agencies to fight terrorism and crime (e.g., Chen et al., 2012; Newell & Marabelli, 2015)[2].

2 Authors have also been critical about the potential undesirable societal effects of deploying data-related technologies (e.g., Clarke, 2016; Newell & Marabelli, 2015; Zuboff, 2015; Loebbecke & Picot, 2015)—this is something I discuss in Chapter 2.

1.4 Research Question

The promise that organizations can achieve social and economic value through informed decision making and the creation of data-driven products and services has largely been backed by supply-side actors such as vendors and consultancy firms (Madsen & Stenheim, 2016). Organizations have enthusiastically jumped on the data-bandwagon with these values in mind. However, anecdotal evidence shows that organizational efforts in this area are not quite living up to the expectations. For example, Gartner (2016) issued a report stating that while organizations are investing in (big) data-related technologies, many of them cannot move past the pilot stage. Likewise, Ransbotham et al. (2016) reported a significant decline in the number of organizations claiming that data and analytics are sources of competitive advantage. More recently, a Gartner survey (2018-2) among 196 organizations has shown that a staggering 91% of respondents indicated that their organizations are not yet able to realize the full potential of data. In other words, there is a gap between what is promised and what is actually realized in practice by organizations.

The aim of this dissertation is to commence bridging this gap. This requires moving beyond the hype around "big data" that is driven by vendors and management gurus (Madsen & Stenheim, 2016), and focus on understanding how organizations actually approach this phenomenon. The overall research question for this dissertation is: How do organizations explore the opportunities of data as strategic resources? By focusing specifically on what tensions organizations face, how data raise opportunities and challenges, and how different actors are involved as organizations aim to realize value from data, I aim to provide insights that can help both scholars and practitioners understand how organizations can effectively leverage data as strategic resources.

1.5 Dissertation Outline

I performed three studies with specific research questions which, taken together, contribute to an understanding of the research question at large. These studies are presented in Chapters 2, 3 and 4. A summary of these chapters and how they are related is presented in Figure 1.1. The fifth chapter provides implications for researchers and practitioners aiming to understand how data may effectively be leveraged as strategic resources.

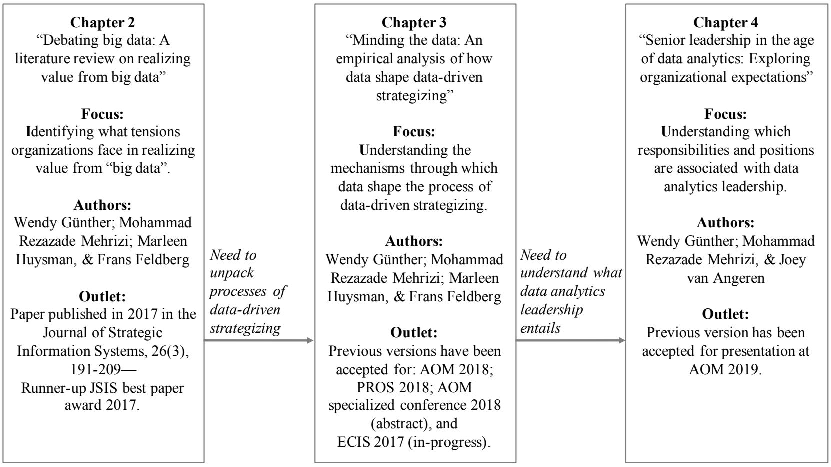

Figure 1.1 Summary of chapters.

1.5.1 First Study: Analyzing the Literature to Identify Tensions

I began my journey by delving into the existing Information Systems literature to find what had thus far been written about how organizations realize social and economic value from data. Literature reviews are especially useful when studying an emerging field, as they allow the researcher to unpack the different theoretical foundations and assumptions that are associated with the topic (Webster & Watson, 2002). My specific aim was to identify *debates* that are central to how organizations realize social and economic value from big data, as this would help me better understand the different tensions that organizations face.

My review of the IS literature consisted of search, selection, analysis and synthesis processes (Boell & Cecez-Kecmanovic, 2015; Jones & Gatrell, 2014; Webster & Watson, 2002). I searched (top) academic journals and proceedings of well-known IS conferences and selected papers that discuss organizational changes, drivers, and actions related to "big data" value realization at different levels of analysis. My search and selection efforts resulted in 67 papers, which I then coded according to a review framework. I analytically abstracted codes around debates at three levels of analysis, that is, the work-practice level (how actors work with data in practice), the organizational level (how organizations organize for big data), and the supra-organizational level (how organizations interact with external stakeholders). For each debate, I articulated two opposing positions and examined the relevant actions that different actors take towards each side, the contextual conditions that shape the debates, and the potential implications for supporting or limiting big data value realization.

The first study is presented in Chapter 2. In this chapter, I present the six debates that resulted from my analysis of the literature and that are central to how organizations realize value from big data. Based on my analysis of the literature, I also critically ask: "what are the potentially unique features of big data that influence data-driven value realization at the work-practice, organizational, and supra-organizational levels?". I present two socio-technological features of big data, "portability" and "interconnectivity", that influence data-driven value realization. I argue that, in practice, organizations need to continuously realign work-practices, organizational models, and stakeholder interests in order to be able to realize social and economic value from data. Finally, I synthesize the findings into an integrated model and provide several avenues for future empirical studies.

1.5.2 Second Study: Learning about Data-Driven Strategizing from a Rich Case

Among other things, I learned from the literature review that there had been few qualitative studies that zoom in on how organizations develop and implement data-driven strategies. In the meantime, organizations were still jumping on the data-bandwagon. This provided me with the opportunity to conduct my second study: an in-depth longitudinal case study of an organization that was triggered to invest in data and data-related technologies by the promises of social and economic value, yet struggled to actually realize such value. The specific research question for this study: "How do data shape the process of data-driven strategizing?" was driven by the observation that studies have largely been pushing the potential role of data to the background (Jones, 2018; Tempini, 2017).

Case studies can be especially useful to study a relatively new phenomenon: "With the rapid pace of change in the information systems field, many new topics emerge each year for which valuable insights can be gained through the use of case research" (Benbasat et al., 1987, p. 370). Essentially, we may first learn how organizations are approaching new phenomena in practice, before we can usefully develop theories and management guidelines (Benbasat et al., 1987). Case studies have also proven to be particularly useful when researchers need to capture phenomena and dynamics inside the context in which they occur (Yin, 1981; Eisenhardt, 1989). Indeed, to answer my research question, I needed to understand what the data in a particular setting looked like and how actors in that setting worked with the data. To understand how data influenced strategic choices and actions *over time*, I adopted a process perspective and focused on the collection and analysis of longitudinal data (Pettigrew, 1990; Pettigrew, 1992; Langley, 2007).

The case concerns a European postal service organization, LogiCo, that had traditionally been delivering mail items to households on behalf of business clients. In 2012, LogiCo's senior management realized that they were facing a shrinking market due to digitization. Driven by the hype around "big data", they thought that they could use data to develop new, data-driven products and services that might help them survive in light of the shrinking market. Over time, however, managers found that the benefits that they had initially hoped for were not being realized in practice. Thus, the case study served as a rich case for understanding the range of challenges and mechanisms that may prevent organizations from successfully leveraging data as strategic resources.

The second study is presented in Chapter 3. In this chapter, I narrate how LogiCo moved into different strategic directions over time, and critically reflect on how data shaped this process. I show that while data can enable strategic exploration, data may also encourage organizations to remain close to the traditional purposes for which they have been collected. Additionally, I show that while organizations may be encouraged to collaborate with external stakeholders because of the nature of the data, such strategic collaborations may also be hindered due to characteristics of the data. I reflect on these tensions and discuss what the findings imply for our understanding of data-driven value realization.

1.5.3 Third Study: Characterizing Senior Leadership in the Age of Data Analytics

One of the lessons learned from the previous study, is that those involved in data-driven strategizing should be able to understand how data can lead them into certain strategic directions. This includes *senior managers*, who may not only be expected to have affinity with the data, but should also be actively involved in, and direct, data-driven initiatives (e.g., Fitzgerald, 2014; Lee et al., 2014; Seddon et al., 2017). Despite a general conviction that "data analytics leadership" by senior managers is important, however, the literature is ambiguous about what such leadership actually entails. In the third study, I explored: "What are the expected responsibilities and positions of data analytics leaders?"

To answer the research question in the third study, I collected and analyzed job ads for senior managers with varying titles from a popular job website, Indeed.com. The analysis of job ads can be "particularly relevant to areas where job roles are increasingly non-traditional and not connected to a well-defined profession" (Harper, 2012, p. 30), which seems very applicable given that I am studying an emerging field. I applied a topic modeling algorithm on the content of the job ads and iterated between the patterns and the job descriptions to explore: 1) What are the data-related responsibilities that characterize data analytics leadership? 2) Which positions will be responsible for providing data analytics leadership? and 3) Which (combinations) of responsibilities are stressed as the most important for "data analytics leaders"?

The third study is presented in Chapter 4. In this chapter, I present four data-related responsibilities, which characterize data analytics leadership in practice: “data infrastructure”, “data control”, “applied analytics”, and “data privacy”. The data suggest that job ads for “data analytics” positions often provide explicit information regarding data-related responsibilities, which implies that organizations do not just signal these terms in the job titles. Conversely, positions with titles pertaining to “information technology” and “digital” are much less likely to be responsible for providing data analytics leadership. Finally, I find that most “data analytics leaders” in our sample will be responsible for ensuring the creation of insights, products, and services from data through analytics, while the privacy implications hereof are only marginally represented and discussed in their job ads. I discuss what these findings imply for our understanding of leadership in the context of data analytics.

1.6 Contribution of this Dissertation

This dissertation contributes to the emerging stream of literature on data-driven value realization, by providing the groundwork for theories on how organizations can effectively leverage data as strategic resources. The studies in this dissertation provide several insights, which advance our understanding of the phenomenon. Specifically, I present different tensions that organizations face as they try to realize value, explain how data can influence strategic choices and actions, and provide a characterization of leadership in the age of data analytics.

As a whole, the dissertation presents a number of learning opportunities for scholars. First, the findings presented in this dissertation concur that, to fully understand how organizations can leverage data as strategic resources, scholars should consider bringing the data to the forefront and more critically examining *what* characteristics of data influence value realization, and *how* this happens (Jones, 2018; Tempini, 2017). Second, the findings affirm that data are extremely dynamic resources, of which the nature and role may continuously change over time as actors work with them in practice (Kallinikos et al., 2013). I also highlight that data are *historically situated*, and point out the relevance of adopting a process lens to understand how organizations may leverage data as strategic resources. Finally, I emphasize that there is an increasing need for research that focuses on how organizations can deal with the (unforeseen) social implications of data analytics.

Practitioners may learn from this thesis that realizing value from data is inherently a cross-level process. Managers need to carefully consider how they may facilitate data collection and analysis at the work-practice level; how they will design organizational models at the organizational level, and how they will deal with stakeholder interests at the supra-organizational level. In order to align efforts at different levels of analysis, different stakeholders—including those at the level of senior management and external stakeholders—may need to be closely involved in data-driven value realization. Additionally, the findings presented in this dissertation suggest that the process of data-driven value realization is often *not* a linear process in which organizations simply collect data, analyze these data, and then engage in value-creating actions. Instead, practitioners may find themselves continuously revisiting the data, their choices, and their actions in a process of exploration. Finally, it is extremely important that managers know *what data* they have and *how* these data may play a role in the process of data-driven value realization. An important part of knowing what data you have also involves knowing where the data *came from*, and how traditional decisions may be inscribed in the data and influence future choices and actions.