

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler and Thomas Plank

“I never guess. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.”

Sir Arthur Conan Doyle,
Author of Sherlock Holmes Stories

Researchers and practitioners agree upon the huge potential of Big Data Analytics (BDA) for firms' competitive advantage. Capitalizing on Big Data (BD) often requires sharing firms' data with their stakeholders in an ecosystem. Sharing data, however, is a double-edged sword, because firms might also risk losing their competitive advantage by doing so. This conceptual paper uses extant literature on data analytics to introduce a comprehensive data sharing strategy framework that helps firms decide which data can be shared in an ecosystem, which should be kept secret, or if additional rules and regulations are required for sharing the data. The framework combines two distinct categorizations of data, and it relates the data categories to firms' strategic competitive advantage by employing the Resource-Based View (RBV). Firstly, the framework is grounded in the stages of the data analytics process and secondly, it distinguishes between operative, strategic and monetizable data, a new categorization introduced by this paper. Depending on the categories of data a company intends to share, the framework recommends five distinct data sharing strategies that help mitigating the risk of losing their competitive advantage.

Introduction

Digital technology and artificial intelligence are fundamentally changing the rules of business competition in markets from an external perspective, as well as the processes of value creation from an internal perspective (Brynjolfsson & McAfee, 2014; Iansiti & Lakhani, 2020). Especially “big data” and “big data analytics” (BDA) create new possibilities for strengthening companies' efficiency and productivity (Aaser et al., 2020; New Vantage Partners, 2020), or for fostering innovativeness and growth options (Aaser et al., 2020; Mariani & Fosso Wamba, 2020; New Vantage Partners, 2020) by enabling new products, processes, business models, or services (Lim et al., 2018; Auh et al., 2021). Big data is seen as a promising resource that has a positive effect on business or societal value (Aaser et al., 2020), competitive advantage, and company

performance (Fosso Wamba et al., 2017; Auh et al., 2021). While the amount of data that is available for firms explodes (Davenport & Bean, 2018), many organizations are still struggling to compete regarding data (Aker et al., 2016; Vidgen, 2017; Urbinati et al., 2019). A recent New Vantage Partners study reported that there has been little to no success for companies over the past years to become data-driven (New Vantage Partners, 2020), and the majority of firms (61%) to date have been unable to turn insights from data into a competitive advantage (Jiang et al., 2021). The gap between leaders and laggards in adopting BDA is growing within and between industries (Diaz et al., 2018; Jiang et al., 2021).

The literature identifies a variety of reasons that hinder firms from turning data into value. Firstly, networking and data sharing are prerequisites for value generating

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

data applications in business ecosystems (Cui et al., 2020). However, sharing data is a double-edged sword because, on the one hand, the data's value increases by sharing it (Lim et al., 2018; Thuermer, 2019) or by gathering and curating the data on sharing platforms (Parra-Moyano et al., 2021). On the other hand, organizations risk losing their source of value and competitive advantage to stakeholders because they run the risk of dependency or exploitation, especially in the longer run. Consequently, these firms are trapped in a data sharing dilemma (Kraemer et al., 2019). It can be concluded that to capitalize on data sharing, firms must first understand the mechanisms of data sharing that include, first, which data they can freely share with their (external) stakeholders, second, which data they need to protect, and, third, what tools and agreements help protect the data without compromising the value that is generated by sharing the data (IMDA & PDPC, 2019).

Secondly, the specific characteristics of data as a resource prove to be a hurdle for turning data into value, because raw data alone are insufficient for the generation of value from it (Gupta & George, 2016; Bumblauskas et al., 2017). Data are an intangible good (IMDA & PDPC, 2019) that is non-exclusive in use (Parra-Moyano et al., 2020). Anyone, or any firm that has access to the data can use it, which makes raw data inadequate for generating a competitive advantage (Parra-Moyano et al., 2020). For capitalizing on data, firms must clean the data, integrate, aggregate, and analyze it in a data analytics process (Jagadish et al., 2014). By doing so (raw) data must first be turned into actionable knowledge (Argyris, 1995), a process that requires both interpretation and integration by humans (Bumblauskas et al., 2017).

Based on an extant review of the literature on data analytics, this conceptual contribution aims at discussing how firms can constructively craft strategies for dealing with the double-edged sword of sharing data in a digital ecosystem. The paper introduces a comprehensive data sharing strategy framework that helps in deciding which company data can be easily shared with a firm's stakeholders without losing possible competitive advantages that can be generated from the data. The framework combines two distinct categorizations of data and relates the data categories to a company's competitive advantage by employing a resource-based view (RBV). Firstly, the framework is grounded in the various stages of the data analytics process (Jagadish et al., 2014). Secondly, it distinguishes

between operative, strategic and monetizable data, a new categorization introduced by this paper. Based on the categories of data a company intends to share, the paper recommends five distinct strategies for sharing data that mitigates the risks of losing a company's advantage.

The second section of the paper presents a summary of the ongoing discussion on big data in the management literature. In the third section, the paper reviews how data and data categories are linked to resources, capabilities, and competitive advantage from a RBV perspective. Also, the contribution introduces a data categorization that is based on the data's strategic value, operative, strategic and monetizable. In section four the paper introduces a data sharing strategy framework, that combines these data categories with the stages in the BDA process and it recommends five distinct strategies for sharing data in an ecosystem. Finally, a discussion on how firms can cope with the double-edged sword of sharing data concludes the contribution.

Big Data and Big Data Analytics

Characteristics of Big Data and Big Data Analytics

The term "big data" refers to large datasets from diverse sources that can be harvested (Urbinati et al., 2019) by using advanced techniques and for supporting various decisions (Chen et al., 2012). Big data analytics (BDA) is characterized as "a holistic approach to manage, process and analyze 5 Vs (i.e., volume, variety, velocity, veracity, and value) in order to create actionable insights for sustained value delivery, measuring performance and establishing competitive advantages" (Fosso Wamba et al., 2015). For turning raw data into value, BDA needs to cover a distinct number of steps within a data analytics process that comprises data acquisition, information extraction and cleaning, data integration, modelling and analysis, interpretation, and deployment (Jagadish et al., 2014).

Big data is characterized by features that distinguish it from other kinds of data (Parra-Moyano et al., 2020). Big data is heterogeneous, often unstructured, or semi-structured, agnostic, haphazard, and trans-semiotic (it is stored in text, image, sound), while other data (in a standard strategy process), in contrast, is homogeneous, structured, purposeful, theory-driven, and mono-semiotic (Constantiou & Kallinikos, 2015). Therefore, "Big Data is different data" (Constantiou & Kallinikos,

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

2015), since it has a different, more difficult quality compared to other kinds of (“traditional”) data. For working with and generating value from the data, a firm requires a complex mix of big data analytics capabilities (Gupta & George, 2016; Akter et al., 2017; Mikalef et al., 2020), meaning, “the ability of a firm to effectively deploy technology and talent to capture, store, and analyze data toward the generation of insight” (Mikalef et al., 2020).

Application of Big Data Analytics Practices in Firms

Despite the huge potential inherent to big data, firms claim that they still find it difficult turning BDA into new businesses or into value (Vidgen et al., 2017; New Vantage Partners, 2020). The potential inherent to the technology to a large extent seemingly cannot be exhausted and many firms cannot generate the competitive advantage or the increase in performance they had expected when using big data and BDA (Hagiu & Wright, 2020).

In fact, turning BDA into value seems to take more than just technology (Storm & Borgman, 2020). Factors inside the company especially must first be aligned to deal with big data, such as having a data-driven organizational culture (Gupta & George, 2016; Upadhyay & Kumar, 2020), a decision-making culture (McAfee & Brynjolfsson, 2012; Vidgen et al., 2017), a data-dominant logic (Kugler, 2020), and technical and managerial skills or roles (Gupta & George, 2016; Davenport & Bean, 2018). Many established processes, objectives, tools, and paradigms do not allow thinking and working with data beyond the established well-known structures (Kugler, 2020). This is especially true when a firm intends to use big data for innovation or strategizing purposes (Constantiou & Kallinikos, 2015).

To achieve the required shift organizations must develop distinct big data analytics capabilities (BDAC, more below), a multi-dimensional construct that covers management capability, technological capability, and talent capability (Akter et al., 2016). Hagiu and Wright (2020) conclude that firms lack data-driven business models and likewise that practitioners generally lack guidance for dealing with data analytics, a key component for addressing differences between experts and laggards (Vidgen et al., 2017).

Data Sharing in Ecosystems

As data is often created “when two or more instances of use interact” (Parra-Moyano et al., 2020) generating

value from data often requires sharing the data in an ecosystem, rather than in a company’s isolated activities (IMDA & PDPC, 2019). “Data sharing” refers to “the sharing of otherwise closed data within or between organizations” (Thuermer et al., 2019). Other options for getting access to data, such as open data (Thuermer et al., 2019) or trading data on the market are difficult or of limited use because data’s characteristics tend to hinder these transactions, and firms consequently tend not to share their data (Parra-Moyano et al., 2020).

In data sharing ecosystems, partner organizations “agree to share data and insights under locally applicable regulations to create new value for all participants” (Jiang et al., 2021). All kinds of organizations can benefit from sharing data, including data holders, innovators, intermediaries, and society as a whole (Thuermer et al., 2019). Data sharing ecosystems go beyond traditional value chains, industries, or data domains and have the potential for generating superior company performance because sharing data improves customer satisfaction (15% annually in the last 2-3 years), productivity and efficiency (14%), and helps reduce costs (11%, Jiang et al., 2020), while shared data enable data-driven innovation (Stalla-Bourdillon et al., 2020).

The Double-Edged Sword of Sharing Data

However, data sharing ecosystems are still in an infant stage (IMDA & PDPC, 2019) and their full potential remains untapped (Jiang et al., 2021), because sharing data is a double-edged sword for the companies involved. On the one hand, the data’s value increases by sharing it (Lim et al., 2018; Jiang et al., 2021), while, on the other hand, organizations risk losing their source of value by granting their partners access to their data. They also run the risk of dependency or exploitation, especially in the longer run (Kraemer et al., 2019).

These firms are trapped in a data sharing dilemma (Kraemer et al., 2019) or, more generally speaking, in a social dilemma caused by data sharing (Linek et al., 2019). Social dilemmas are characterized when selfish, non-cooperative behavior is deemed more beneficial to individual parties involved. Yet if all parties involved behave in a non-cooperative way, they all would receive less payoff than if everyone cooperated (Linek et al., 2019). This risk runs especially high for small or young organizations that are sharing data with large platforms (Kraemer et al., 2019). The data sharing firms face a trade-off between positive short-term effects of sharing

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

and challenging long-term strategic effects (Kraemer et al., 2019). Companies intending to get involved in these ecosystems still require guidance to help them cope with the challenges of sharing data, such as understanding the mechanisms of sharing data, ensuring compliance to regulations, and establishing mutual trust (IMDA & PDPC, 2019).

Big Data Categories from a Resource-Based View

Big Data Categories

The management literature on big data discusses a broad variety of data categories such as characterizing data as a resource or capability (Gupta & George, 2016; Bumblauskas et al., 2017; Mariani & Fosso Wamba, 2020), steps in the data analytics process (Jagadish et al., 2014), structured, semi-structured, and unstructured data (Praveen & Chandra, 2017), and the data related dimensions of volume, variety, velocity, and value (Aker et al., 2016; Fosso Wamba et al., 2017). The large number of data taxonomies available indicates that there is no one-size fits all solution to categorizing data, but rather depends on the context organizations need to define their own data categories (IMDA & PDPC, 2019). Also, none of these taxonomies alone is enough to determine the data's value or how it contributes to a competitive advantage. Similarly, Bumblauskas et al. (2017) stated, "the size, scope and scale of data are difficult to limit in defining Big Data, the definition of Big Data must revolve around the analysis of the data rather than the actual size of the data or spreadsheet (i.e. large data sets or databases)". Given these categories, it remains unclear if the data should be limited to focal organization, or if data sharing in an ecosystem is an option. Against this background, the current paper introduces a data taxonomy according to "how strategic the data is to the organization" (IMDA & PDPC, 2019), by the data's potential for generating a competitive advantage, and how it distinguishes between operative, strategic, and monetizable data.

In what follows, the paper introduces a comprehensive framework that builds upon two categorizations of data. First, it is grounded in stages of the BDA process and whether the data can be classified as resources or capabilities. Second, it is based on the data's strategic value regarding whether it is operative, strategic, and monetizable. Both data categories are linked to the data's potential for generating competitive business advantage. Depending on the data available, the paper presents five distinct strategies for data sharing.

Big Data and Big Data Analytics as Resources and Capabilities

In line with Gupta and George (2016), this paper argues that the resource-based view (RBV) links an organization's resources and capabilities (independent variables) with organizational competitive advantage and performance (dependent variables) (Amit & Schoemaker, 1993). Resources characterize "stocks of available factors that are owned or controlled by the firm" (Amit & Schoemaker, 1993). Organizational capabilities aim at connecting and exploiting organizational resources, meaning "the ability of an organization to perform a coordinated set of tasks, utilizing organizational resources, for the purpose of achieving a particular end result" (Helfat & Peteraf, 2003). What resources an organization has and how it combines and uses these resources with its capabilities, directly influences a firm's performance. Yet only those resources and capabilities have the potential to create sustainable advantage that are valuable, rare, difficult to imitate, and without substitutes (Barney, 1991).

Gupta and George (2016) classified (raw) data and their merging as a tangible resource because they are non-exclusive in use and available to many firms in the market (Parra-Moyano et al., 2020). Following Bumblauskas et al. (2017), raw data alone are of no value or only of little value to a certain company, given that they must first be transferred into actionable knowledge (Davenport & Prusack, 1998) that enables people to act or to decide: "[Raw] data is a set of discrete, objective facts about events ... [but] data by itself has little relevance or purpose" (Davenport & Prusack, 1998). Raw data, therefore, does not suit Barney's (1991) four criteria, and it can hardly be a source of competitive advantage alone (Gupta & George, 2016).

Big data analytics capability (BDAC), in contrast, is more complex than raw data and marks a company's ability "to effectively deploy technology and talent to capture, store, and analyze data, toward the generation of insight" (Mikalef et al., 2020). While BDAC adds meaning to raw data, it has the potential to turn data into what Davenport and Prusack (1998) term "information and knowledge", however, this step can only be accomplished by human beings (Gupta & George, 2016), that is, not just by technical means. BDAC, therefore, is firm-specific, and has the potential to be valuable, rare, difficult to imitate, and without substitutes, as Barney (1991) suggests.

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

Table 1. Characteristics of Big Data and Big Data Analytics.

Unit of analysis	Characterization (Davenport & Prusack, 1998)	Step in Big Data Analytics (BDA) process (Jagadish et al., 2014)
<i>Data (tangible / resource)</i>	Data is a set of discrete, objective facts about events.	Data acquisition
<i>Information</i>	Information is data to which meaning has been added, such as by means of contextualizing, categorizing, calculating, correcting, or condensing the data*.	Information extraction & cleaning Data integration Modelling
<i>Knowledge (intangible / capability)</i>	Knowledge is a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knowers.	Analysis Interpretation Deployment

Big Data Analytics Process

For turning raw data into knowledge and into value, raw data must go through a multi-step analytics process that covers data acquisition, information extraction and cleaning, data integration, modelling and analysis, interpretation, and deployment (Jagadish et al., 2014). Although these steps can partly be automated, the complex steps of data analysis, interpretation, and deployment especially depend upon human beings to extract or add meaning to the data (Jagadish et al., 2014; Bumblauskas et al., 2017). The software-generated results must be understood, questioned, or summarized as working hypotheses, all of which requires human cognition (Constantiou & Kallinikos, 2015).

It can be concluded that the biggest potential for companies that wish to turn data into competitive advantage lies in the more advanced steps of the data analytics process (Jagadish et al., 2014) that require BDAC. Meanwhile, less potential resides in the initial steps of the process grounded in informational resources and raw data. For a summary, see Table 1.

*Contextualized: we know for what purpose the data was gathered; categorized: we know the units of analysis or key components of the data; calculated: the data may have been analysed mathematically or statistically; corrected: errors have been removed from the data; condensed: the data may have been summarized in a more concise form (Davenport & Prusack, 1998).

Taking a RBV has been widely used to approach and explain causal relationships between big data, BDA, and competitive advantage (Akter et al., 2016; Gupta & George, 2016; Fosso Wamba et al., 2017; Mikalef et al., 2020). While the RBV is an established approach from an empirical scientific point of view, we conclude that it largely remains on an abstract or theoretical level, and typically lacks a comprehensive approach that helps get an overview of the available data's potential to generate competitive advantage.

Operative, Strategic and Monetizable Data

This paper introduces another categorization of data that is based on the data's strategic value. It distinguishes between operative, strategic, and monetizable data because these categories give information on how an organization uses or intends to use a certain data set in the shorter or longer run. While *operative data* are necessary to run daily business, *strategic data* can be used for innovation activities, while *monetizable data* are of little use for the focal company itself, but are rather of great use for external stakeholders.

Operative data are used to run current business. Data is used for efficiency increases, such as for controlling or (predictive) maintenance purposes, and for digital twins. These improvements can lead to short-term cost advantages in competition, while the advantages gained from the data might erode over time. Competitors can

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

Table 2. Examples of operative, strategic, monetizable data

operative data	strategic data	monetizable data
(predictive) maintenance	customer insight	raw data
data supply chain	consumer trends	aggregated data
controlling	product, process, service	data insights
digital twin (product, manufacturing)	innovation	
monitoring	business model innovation	
product, process, service improvements	management innovation	
efficiency increase	new business opportunities	additional yield

also use similar data in a similar way. Therefore, data itself cannot provide the potential for generating a competitive advantage (Kraemer et al., 2019). It is rather likely that using data creates a new standard in competition that is already or will later be used by many players in global markets.

Strategic data are used for generating new business opportunities, including innovative products services, processes, or business models, that might, for instance, build upon available consumer data. Strategic data has the potential to enable new possibilities for a firm's future success. New business opportunities link prior knowledge and solutions to unknown insights, and are more complex than operative solutions. Strategic data have the potential to serve as the origin of gaining competitive advantages.

Monetizable data are data than can be sold to stakeholders, such as, for instance, data that was generated as a by-product of other activities, and that is of no or of little use for the company. Such data can be used for generating additional yield for an organization. Similar to operational data, monetizable data are rather unlikely to serve as the origin of a competitive advantage (for examples see also Table 2).

Data Sharing Strategy Framework

Data Sharing Strategies

In the following section, we correlate the data categories operative, strategic, and monetizable with steps in the BDA process (Jagadish et al., 2014). As was demonstrated above, the highest value for a company resides in the more complex final stages in this process that require profound data analytics capabilities (data

analysis, interpretation, and deployment; Bumblauskas et al., 2017). These activities are necessarily linked to the interpretations and experiences which people add to making sense of the data (Bumblauskas et al., 2017), which is difficult to imitate or substitute. Largely unprocessed raw data are particularly valuable only if no other company has comparable data available. Therefore, raw data alone are of little or no strategic value to organizations. However, there is no guarantee, but rather only a probability that some kind of competitive advantage can be generated from the data.

From the proceeding discussion, it can be concluded that, depending on the type of data available, different strategies are available for how to deal with the data. Correlating the steps in the big data analytics process with the potential uses and strategic value of the data leads to five different strategies, depending on how well they are suited to generate a competitive advantage. These strategies will be outlined below, along with a brief illustration of each (see Figure 3). Strategies (1) and (2) are extreme cases in which data should or should not be shared openly at all with a company's stakeholders:

Strategy (1) deals with big data that classifies as strategic and that has been analyzed or interpreted. It therefore represents (actionable) knowledge that can be of great value to a company. The analysis and/or interpretation of the data strongly depends on the company's BDAC. These types of data have a high probability of leading to a competitive advantage. Companies should clarify the possible gains and risks of sharing these data. The safest way to cope with the potential risk is not sharing it at all.

However, if potential gains can outweigh the potential risks, then a company should use clear mechanisms to

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

Table 3. Data sharing strategy framework

<i>knowledge</i> deployment interpretation Analysis	(5) potential source of competitive advantage (sell / share the data, risk mitigation)	(4) potential source of competitive advantage (sharing is an option, risk mitigation)	(1) probable source of competitive advantage (no sharing)
<i>information</i> modeling data integration information extraction & cleaning	(2) no potential source of competitive advantage (sell / share the data)	(3) potential source of competitive standard (sharing is an option, risk mitigation)	(4) potential source of competitive advantage (sharing is an option, risk mitigation)
<i>raw data</i> data acquisition	(2) no potential source of competitive advantage (sell / share the data)	(3) potential source of competitive standard (sharing is an option, risk mitigation)	(4) potential source of competitive advantage (sharing is an option, risk mitigation)
<i>step in big data analytics process</i> <i>use of big data</i>	monetizable data (additional yield)	operative data (efficiency increase)	strategic data (new business opportunities & innovation)

mitigate the risks. A brief example illustrates the strategy:

Most major car manufacturers (for example, BMW, Audi, Ford) equip their new vehicle models with sensors that enable them to provide telematics data (raw data) that is collected in the company's data center. The telematics-enabled vehicles generate a wide range of data, including condition data (for example, mileage), usage data (for example, heavy acceleration), or event data (for example, power interruption, service call). If selected data is aggregated and analyzed, it can be the starting point for the vehicle manufacturer's new services, products, or business models. This is the case, for example, if a company wants to offer its customers a predictive service model that can use the combination of data to predict when a vehicle repair is very likely to be necessary. The aggregated and analyzed data then has strategic value for the company and is not shared.

Strategy (2) combines monetizable data that is of no or little use for a company at the initial stage(s) in the big data analytics process. The data makes no use or only to a very small degree uses a firm's BDAC. Therefore, the probability that the data could lead to a competitive advantage is low. However, selling the data can lead to generating additional yield from the data (once or repeatedly), or to receiving some extra information for

the data.

In this case the car manufacturer collects telematics data (raw data) that provides information about the use of the cars' shock absorbers. This data has already been collected for several years. Additionally, data is also collected that allows drawing conclusions about the condition of the roads cars are driving on, especially with respect to potholes, which place particularly high stress on shock absorbers. The car manufacturer has no use for the road condition data. However, that data may be of interest to a city, municipality, or country for the purpose of infrastructure maintenance, otherwise the government would have to pay the price of collecting this data by itself. The car manufacturer can sell the raw data to the city or country or share it for a fee. Alternatively, the vehicle manufacturer can provide the data free of charge to the government, but get back an aggregated view (more valuable data) from the government that is created from data provided by all manufacturers in return.

Strategy (3) deals with operative raw data or information that is helpful for efficiently running a daily business. Similar data is available or can be generated without the need for concise analytics capabilities by many companies. Therefore, the probability for generating a competitive advantage is rather low. However, because

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

this type of data is easily available, its use can rapidly lead to a competitive (or industry) standard. This data is then more of a prerequisite to compete within an industry or between ecosystems.

The car manufacturer collects data on vehicle use, which provides information on the wear and tear of the vehicle's parts (for example, tires, battery, etc.), which the car manufacturer does not produce itself, but which it obtains from suppliers. This data can be interesting for suppliers (customers) because it helps to use these wearable parts more efficiently, for example if wearing of the tires depends more on climate or the driving behaviour of the driver. This data can be shared with suppliers as raw data, while the vehicle manufacturer employs additional means for risk mitigation.

Strategy (4) either consists of strategic raw data and information that can generate new business opportunities, or of operative data that has already been turned into actionable knowledge, and therefore reflects a potential source of competitive advantage. Firms should decide case-by-case if potential gains that can be achieved through sharing outweigh the risks of not doing so, like strategy 3.

The car manufacturer collects raw data on the driving behaviour of vehicle owners and on vehicle use, which provides information on the accident behaviour (probability of an accident) of drivers, and optionally the vehicle manufacturer aggregates and analyzes the data. These data form the basis for new business models for insurance companies that depend on driving behaviour and frequency of use reports. Such data would otherwise have to be collected separately by the insurance company. The data can be shared with or sold to the insurance company (as long as the drivers give their informed consent according to the regional legal standards).

Strategy (5) refers to monetizable data that has been transferred into knowledge by means of combining data analytics practices and capabilities. While these data are of no or little value to focal firms, the possibility still exists for generating competitive advantage given that the analyses can be valuable, rare, and difficult to imitate or substitute (Barney, 1991). Sharing or selling the data, therefore can be an option, but, again, the company should consider the appropriate risk mitigating activities.

Strategy 5 is like strategy 2, albeit with data that have already been aggregated, analyzed, and processed, instead of using raw data. These data can serve as the basis for a new business model.

It can be concluded that the highest value for firms resides in strategic data of all kinds, and especially in those data that have already been turned into actionable knowledge by means of analytics in combination with BDA capabilities. But also, operative or monetizable data that has been analyzed has a similar potential for value creation, and it should be protected or subject to the mitigation of potential risks.

Literature on Mitigating the Risks of Data sharing

The data analytics framework presented in this paper gives an overview of strategies for sharing data in a digital ecosystem, but it gives minimal information on concrete activities about how firms can mitigate potential risks that may arise. The scientific literature only starts to discuss a variety of measures that companies can take to mitigate risks, yet without classifying these activities and on a rather broad, unspecific level.

Some authors suggest using data trusts (Protection Information Management, 2018; Stalla-Bourdillon et al., 2021), making data sharing agreements (IMDA & PDPC, 2019) or contracts (Thuermer et al., 2019) when confronted with the risk of sharing data. While no prescribed format currently exists for such agreements, these companies and other sharing organizations should agree upon key issues, such as data confidentiality, the allocation of liability for contract breaches (IMDA & PDPC, 2019), restrictions to permitted data usage, and clarifications about who owns any intellectual property outcome of the shared data (Thuermer et al., 2019). In any case, trust between the sharing partners seems to play a crucial role for mitigating the risks of data sharing in digital ecosystems, and it can be strengthened by following the principles of fairness and ethics, transparency, security, and data integrity (IMDA & PDPC, 2019).

Kraemer et al. (2019) by referring to partnerships with large online platforms, suggested seeking data sharing partners from complementary markets or strengthening differentiation between competitors through sharing partnerships. Other literature refers to technical issues

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

for mitigating the risks of sharing data, such as applying algorithms to data only where the data originally is stored so that raw data never leaves its repository, applying open algorithms so that experts can judge an algorithm's safety, or keeping data always protected in an encrypted state (Parra-Mayano et al., 2020). However, the discussion of how to cope with shared data is only in its beginning stages, and further research is required to better understand the appropriateness of the suggested patterns of risk mitigation in sharing partnerships, as well as the processes to do so.

Discussion and Conclusions

This paper aimed at, first, presenting a data sharing strategy framework that relates different types of data to decisions whether an organization should share their data in a digital ecosystem or not. Secondly, the paper introduced a comprehensive classification of big data that links data to competitive advantage, and distinguishes between operative, strategic, and monetizable data that correlates with steps in the BDA process. The paper adds value to both the scientific community and to companies that wish to share data in their ecosystems.

Practitioners can profit from the data framework by getting an overview of various data categories, and of the different strategies for sharing data while mitigating the risks of losing their competitive advantage. To *scientists*, the framework conceptually links the new topic of data sharing to well-established theoretical concepts such as capabilities and resources. However, data sharing in ecosystems is still a new topic that is only starting to be discussed in the scientific literature, and companies are still in search of answers to many questions about sharing their business data. Therefore, some issues can be identified that remain open to future studies. First, as the paper was developed conceptually, the findings of this contribution should be further verified by using empirical evidence.

Second, the paper assumes that data can clearly be classified by their strategic relevance and using a data analytics process. However, for companies these classifications might not always be clear, because on the one hand, firms might lack some pieces of information that would help them to classify their data as operative, strategic, or monetizable. Whenever firms get access to new pieces of information or to additional new data that can be combined with prior findings, the data's strategic

value can be subject to changes. On the other hand, what value a set of data has, differs between firms and between the context in which the data is used. New partnerships in a digital ecosystem or new possibilities to which the data can be applied, therefore, have the potential to also change how the data can be used and, finally, classified. Firms cannot always clearly determine what they will work on in the future. This is also a reason why numerous firms are collecting huge amounts of unstructured data, although they do not yet have a concrete purpose for using the data.

Thirdly, future research could refer more in detail to activities that firms can take to mitigate the risks of sharing data. The lack of a detailed overview persists of concrete measures and of a discussion of which activities are suited best for which data sharing situations. Not all options are open to all firms, due to constraints, such as customers that clearly define what their suppliers are allowed to do with the data (and often they are not allowed to do anything with the data at all). Although many of these measures that mitigate the risks of sharing data seem to be obvious at first sight, their application in a concrete situation of data sharing leaves many questions open. Such as, for instance, how can we clearly determine the value of data for a certain company? How can we estimate all possible risks and benefits of sharing data? How can we overcome internal or external hurdles for sharing data? What could a data sharing contract between digital ecosystem partners look like that constructively deals with the intangible and changing nature and value of big data?

Finally, and closely related to the proceeding issue, it also became clear to us, that for firms it might be not enough anymore to consider strategy and competitive advantages on the level of a single firm only. The more firms start to become part of broader ecosystems, the more it will be necessary to also take into account the perspective of the entire system, also when it comes to competitive advantages. The challenge will be to balance advantages on the firm level with those on the ecosystem level, while being aware of the potential contradictions or trade-offs that may arise in such situations. Companies should also reflect on the purposeful and comprehensive tools and approaches available for how to deal with possible contradicting goals on the firm and ecosystem levels, especially when sharing their data.

The Gartner Group (Goasduff, 2021) suggests, for

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

instance, a “must share data unless” policy. Firms, then, are supposed to share their data unless there is a valid reason why this should not be the case. Open data, consequently, would become a new standard instead of proprietary data and information, which is today’s standard. Open source software projects already demonstrate how such a change in the overall data regime can function. These projects refer to “copyleft” instead of “copyright” (Stallman, 2007). This implies, however, that firms must also change their mindsets and organizational cultures when sharing their data.

Acknowledgments

This study was supported by the grant “ABH097 Data Sharing Framework” within the framework of the Interreg VI-programme “Alpenrhein-Bodensee-Hochrhein” (DE/AT/CH/LI), with funds provided by the European Regional Development Fund (ERDF) and the Swiss Confederation. The funders had no role in the study’s design, data collection and analysis, decisions to publish, or preparation of the manuscript.

References

- Aaser, M., Kanagasabai, K., Roth, M., & Tavakoli, A. 2020. Four Ways to Accelerate the Creation of Data Ecosystems. *McKinsey Analytics*. DOI: <https://doi.org/10.1049/ic.2014.0031>
- Akter, S., Fosso W.S., Gunasekaran, A., Dubey, R., & Childe, S.J. 2016. How to Improve Firm Performance using Big Data Analytics Capability and Business Strategy Alignment. *International Journal Production Economics*, 182: 113-131. DOI: <http://dx.doi.org/10.1016/j.ijpe.2016.08.018>
- Amit, R., & Schoemaker, P.J.H. 1993. Strategic Assets and organizational Rent. *Strategic Management Journal*, 14(1): 33-46. DOI: <https://doi.org/10.1002/smj.4250140105>
- Argyris, C. 1995. Action Science and Organizational Learning. *Journal of Managerial Psychology*, 10(6): 20. DOI: <https://doi.org/10.1108/02683949510093849>
- Auh, S., Menguc, B., Sainam, P., & Jung, Y.S. 2021. The Missing Link between Analytics Readiness and Service Firm Performance. *The Service Industries Journal*. Ahead-of-Print: 1-30. DOI: <https://doi.org/10.1080/02642069.2021.1998461>
- Barney, J. 1991. Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1): 99-120. DOI: <https://doi.org/10.1177/014920639101700108>
- Brynjolfsson, E., & McAfee, A. 2014. *The Second Machine Age. Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York: Norton.
- Bumblauskas, D., Nold, H., Bumblauskas, P., & Igou, A. 2017. Big Data Analytics: Transforming Data into Action. *Business Process Management Journal*, 23(3): 703-720. DOI: <https://doi.org/10.1108/BPMJ-03-2016-0056>
- Chen, H., Chiang, R.H., & Storey, V.C. 2012. Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4): 1165-1188. DOI: <https://doi.org/10.2307/41703503>
- Constantiou, I.D., & Kallinikos, J. 2015. New Games, New Rules: Big Data and the changing Context of Strategy. *Journal of Information Technology*, 30: 44-57. DOI: <https://doi.org/10.1057/jit.2014.17>
- Cui, Y., Kar, S., & Chan, K.C. 2020. Manufacturing Big Data Ecosystems: A systematic Literature Review. *Robotics and Computer Integrated Manufacturing*. 62: 101861. DOI: <https://doi.org/10.1016/j.rcim.2019.101861>
- Davenport, T., & Prusack, L. 1998. *Working Knowledge: How Organizations Manage what they Know*. Boston: Harvard Business School Press.
- Diaz, A., Rowshankish, K., & Saleh, T. 2018. Why Data Culture matters. *McKinsey Quarterly*, September.
- Fosso Wamba, S.F., Akter, S., Edwards, A., Chopin, G., & Guanrou, D. 2015. How 'Big Data' can make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study. *International Journal of Production Economics*, 165: 234-246. DOI: <https://doi.org/10.1016/j.ijpe.2014.12.031>
- Fosso Wamba, S.F., Gunasekaran, A., Akter, S., Ren, S.J., Dubey, R., & Childe, S.J. 2017. Big Data Analytics and Firm Performance: Effects of Dynamic Capabilities. *Journal of Business Research*, 70: 356-365. DOI: <https://doi.org/10.1016/j.jbusres.2016.08.009>
- Goasduff, L. 2021. Data Sharing Is a Business Necessity to Accelerate Digital Business. *The Gartner Group*. <https://www.gartner.com/smarterwithgartner/data-sharing-is-a-business-necessity-to-accelerate-digital-business> [Accessed 12 May 2021].
- Gupta, M., & George, J.F. 2016. Toward the Development of a Big Data Analytics Capability. *Information and Management*, 53(8): 1049-1064. DOI: <https://doi.org/10.1016/j.im.2016.07.004>
- Hagi, A., & Wright, J. 2020. When Data creates Competitive Advantage. *Harvard Business Review*, 98(1): 94-101.
- Helfat, C.E., & Peteraf, M. 2003. The Dynamic Resource-Based View: Capability Lifecycles. *Strategic Management Journal*, 24(10): 997-1010. DOI: <https://doi.org/10.1002/smj.332>
- Iansiti, M., & Lakhani, K.R. 2020. *Competing in the Age of AI: Strategy and Leadership when Algorithms and Networks run the World*. Boston: Harvard Business Review Press.

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

- IMDA Infocomm Media Development Authority of Singapore, & PDPC Personal Data Protection Commission. 2019. *Trusted Data Sharing Framework*. Singapore.
<https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Trusted-Data-Sharing-Framework.pdf>
 [Accessed 22 January 2022].
- Jagadish, H.V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J.M., Ramakrishnan, R., & Shahabi, C. 2014. Big Data and its Technical Challenges. *Communications of the ACM*, 57(7): 86-94. DOI: <https://doi.org/10.1145/2611567>
- Jiang, Z., Thieulent, A.L., Jones, S., Perhirin, V., Baerd, M.C., Shagrithaya, P., Cecconi, G., Isaac-Dognin, L., Buvat, J., Khadikar, A., Khemka, Y., & Nath, S. 2021. *Data Sharing Masters. How Smart Organizations Use Data Ecosystems to Gain an Unbeatable Competitive Edge*. Capgemini Research Institute.
- Kraemer, J., Schnurr, D., & Wohlfarth, M. 2019. Trapped in the Data-sharing Dilemma. *MIT Sloan Management Review*, 60(2): 22-23.
- Kugler, P. 2020. Approaching a Data-dominant Logic. *Technology Innovation Management Review*, 10(10): 16-28. DOI: <http://doi.org/10.22215/timreview/1393>
- Lim, C., Kim, K.H., Kim, M.J., Heo, J.Y., Kim, K.J., & Maglio, P.P. 2018. From Data to Value: A Nine-Factor Framework for Data-based Value Creation in information-intensive Services. *International Journal of Information Management*, 39: 121-135. DOI: <https://doi.org/10.1016/j.ijinformgt.2017.12.007>
- Linek, S.B., Fechner, B., Friesike, S., & Hebing, M. 2017. Data Sharing as Social Dilemma: Influence of the Researcher's Personality. *PLoS ONE*, 12(8): e0183216. DOI: <https://doi.org/10.1371/journal.pone.0183216>
- Mariani, M.M., & Fosso Wamba, S. 2020. Exploring how Consumer Companies Innovate in the digital Age: The Role of Big Data Analytics Companies. *Journal of Business Research*, 121: 338-352. DOI: <https://doi.org/10.1016/j.jbusres.2020.09.012>
- McAfee, A., & Brynjolfsson, E. 2012. Big Data: The Management Revolution. *Harvard Business Review*, 90(10): 60-68.
- Mikalef, P., Krogstie, J., Pappas, I.O., & Pavlou, P. 2020. Exploring the Relationship between Big Data Analytics Capability and Competitive Performance: The mediating Roles of Dynamic and Operational Capabilities. *Information & Management*, 57(2): 103169. DOI: <https://doi.org/10.1016/j.im.2019.05.004>
- New Vantage Partners NVP 2020. *Big Data and AI Executive Survey 2020: Executive Summary of Findings Research Report*. Boston et al.
- Parra-Moyano, J., Schmedders, K., & Pentland, A. 2020. What Managers need to know about Data Exchanges. *MIT Sloan Management Review*, 61(4): 39-44.
- Praveen, S., & Chandra, U. 2017. Influence of Structured, Semi-structured, Unstructured Data on various Data Models. *International Journal of Scientific & Engineering*, 8(12): 67-69.
- Protection Information Management PIM. 2018. *Framework for Data Sharing in Practice*. <http://pim.guide/wp-content/uploads/2018/05/Framework-for-Data-Sharing-in-Practice.pdf>. [Accessed 24 November 2021].
- Stalla-Bourdillon, S., Thuermer, G., Walker, J., Carmichael, L., & Simperl, E. 2020. Data Protection by Design: Building the Foundations of trustworthy Data Sharing. *Data & Policy*, 2(4): 1-10. DOI: [10.1017/dap.2020.1](https://doi.org/10.1017/dap.2020.1)
- Stallman, R. 2007. GNU General Public License. *The GNU Project*. <https://www.gnu.org/licenses/gpl-3.0.html>. [Accessed 28 January 2022]
- Storm, M., & Borgman, H.P. 2020. Understanding Challenges and Success Factors in creating a Data-driven Culture. *Proceedings of the 53rd Hawaiian International Conference on System Sciences HICSS*.
- Thuermer, G., Walker, J., & Simperl, E. 2019. Data sharing Toolkit. Lessons learned, Resources and Recommendations for sharing Data. *Data Innovation Program*, funded by the European Commission. <http://eprints.soton.ac.uk/id/eprint/436050>. [Accessed 25 November 2021].
- Upadhyay, P., & Kumar, A. 2020. The Intermediating Role of Organizational Culture and Internal Analytical Knowledge between the Capability of Big Data Analytics and a Firm's Performance. *International Journal of Information Management*, 52: 102100. DOI: <https://doi.org/10.1016/j.ijinformgt.2020.102100>
- Vidgen, R., Shaw, S., & Grant, D.B. 2017. Management Challenges of creating Value from Business Analytics. *European Journal of Operational Research*, 261: 626-639. DOI: <https://doi.org/10.1016/j.ejor.2017.02.023>

Coping with the Double-Edged Sword of Data Sharing in Ecosystems

Petra Kugler & Thomas Plank

About the Authors

Petra Kugler is a Professor of Strategy and Management at OST - Eastern Switzerland University of Applied Sciences, Switzerland. Her research focuses on the intersection of innovation, strategy, and management, and how firms can generate and protect sustainable competitive advantages in turbulent times. She obtained her PhD from the University of St. Gallen (HSG), has also worked in advertising, and has gained international academic experience through various scholarships, among others a Swiss National Science Foundation Grant for a research year at the University of California, Berkeley.

Thomas Plank is CEO & Co-Founder of Tributech Solutions GmbH headquartered in Linz, Austria. He focuses on the transformation of data to corporate assets and how these assets are managed in the future. He is a Certified Risk Manager and holds a Masters' degree in Controlling, Accounting, and Financial Management of the University of Applied Sciences Upper Austria.

Citation: Kugler, P. & Plank, T. 2021. Coping with the Double-Edged Sword of Data Sharing in Ecosystems. *Technology Innovation Management*, 11 (11/12): 5-16. <http://doi.org/10.22215/timreview/1470>

Keywords: Big Data Analytics Capabilities, Resource-Based View, data sharing, ecosystem, competitive advantage

