A Work Project presented as part of the requirements for the Award of a Master Degree in Management from NOVA – School of Business and Economics

BIG DATA: THE PATH TO DATA MONETIZATION

CHRISTIAN PFEIFER, 30720

A Project carried out in the Master in Management Program under the supervision of:

Professor Fernando Bação

December 29, 2018

Abstract

Big Data: The path to data monetization

The path to monetizing data is a challenging task for many organizations. Not only the technical complexity and the most often needed organizational changes, but also high initial investment costs without certain outcome make this endeavor highly risky. As a result, the number of organizations that have successfully walked the path to monetizing data remains scarce. To address this issue, this paper aims to shed light on the limited understanding of *how* organizations can unlock the value of data and monetize it by conducting an in-depth systematic literature review.

Keywords:

Big Data
Data business model
Data monetization
Value

Table of Contents

1. Introduction

Big Data has gained enormous momentum over the last few years. According to Google Trends, the search interest is over six times higher than in January 2012, providing over 3.9 billion results in October 2018 (Google Trends Explore, 2018). Yet, looking beyond the buzz, the concept of Big Data is not new. It has been around for many years, however, various recent developments in this field, such as more sophisticated tools and analytic capabilities started to unfold the true potential of Big Data (Grable & Lyons, 2018). Consequently, it grabbed the attention of academics and decision makers in government, science, and numerous industries. Many of them fuelled by the hope to monetize data and generate highly valuable insights hidden in the large data sets and partly by the fear to lose competitiveness.

The potential of Big Data is evident throughout today's academic and practitioner literature. Many scholars, practitioners and industry leaders are convinced that it has the potential to transform many fields through its potential to generate valuable insights and to enhance decision-making processes. For instance, Eric Schmidt, CEO of Alphabet, is convinced that "the future of […] business is Big Data and machine learning applied to the business opportunities, customer challenges,…" (Schmidt, 2018, p. 1) In addition, McAfee & Brynjolfsson (2012) argue that decisions and predictions can be made more precisely than ever before, while companies embracing data-driven decisions and deriving insights from Big Data, were, on average, 6% more competitive and 5% more productive than their competitors. Furthermore, Gartner declared Big Data as one of the Top 10 trends in technology in both of its major reports in 2013: "Top 10 Strategic Technology Trends For 2013" and "Top 10 Critical Tech Trends For The Next Five Years". (Zhang & Chen, 2014) Just recently, in September 2018, the importance of Big Data was again accentuated by Accenture in its annual "Accenture

Technology Vision 2018" report as it finds all five major technology trends being data-centric – namely Citizen AI, Extended Reality, Data Veracity, Frictionless Business, and the Internet of Things. These trends are expected to have a huge impact on businesses and society and are forecasted to unfold in the next three years. (Accenture, 2018)

Business model innovation through Big Data is another example that highlights the relevance and impact of Big Data in the business world. It changes the way value is created, delivered and appropriated in modern companies today. Modern companies leverage internal and external data to develop new business models. As mentioned by Dr. Michael Gryseels, Senior Partner at McKinsey & Company, during the "Digital Thailand Big Bang" conference, the world's four largest Unicorns in terms of market capitalization namely Uber, Airbnb, DiDi, and Ant Financials have one thing in common – a data-based business model. They are advanced data-driven companies that disrupted the transportation, hospitality, and payment industry. (Gryseels, 2017)

Knowing the potential of Big Data, public and private organizations from various fields, including business, health care, research, and government, have a strong interest in accumulating data and harnessing its power to differentiate themselves from the competition. However, despite the recent excitement and interest, many organizations have a limited understanding of how to monetize their data and struggle to successfully implement Big Data initiatives (Mikalef, Pappas, Krogstie, & Giannakos, 2016). As a matter of fact, in 2015, Gartner Inc. (2018) estimated that 60% of Big Data projects in 2017 will be abandoned due to the failure to complete the piloting and experimentation stage. Over the past few years, scientific research and corporate investigation on Big Data has received increasing attention. Yet, important contributions have primarily focused on the technical mechanisms, processes and analytical tools of Big Data, such as the highly-cited publication on emerging business intelligence and analytics from Chen, Chiang, & Storey (2012). Contributions on how organizations can assess the business value and strategically monetize their data remain scarce.

Hence, drawing on a systematic review, the purpose of this paper is two-fold. First, this paper aims to help broaden the understanding of Big Data by providing a detailed description. Second, to the best of my knowledge, this is the first study that is to bring forth a synthesis of the latest findings from research in academia and industry reports on Big Data and data monetization strategies. Thus, providing deeper insights into the key challenge in Big Data: "How can firms unlock the value and monetize Big Data?" In doing so, this paper helps business leaders and academics to understand the transformative potential of Big Data as well as the effect that it will undoubtedly have on future business models. For the present purpose of this paper, I focus only on high-level strategic aspects as investigating technical aspects of Big Data monetization would exceed the scope of this study.

This paper is organized as follows. After the introduction (Section 1), I begin the paper in Section 2 by developing a definition of Big Data and data monetization based on existent literature. I emphasize the fact that "Big" is just one of several dimensions of Big Data. Section 2 is followed by an introduction of the research methodology (Section 3). Section 4 investigates the different potential types of data monetization and syntheses the latest findings on the path to data monetization. Section 5 concludes the paper and discusses research limitations and suggestions for future research.

2. Defining Big Data and data monetization

As mentioned in the introduction, Big Data has attracted tremendous attention and triggered many scientific research initiatives and corporate investigations over the last few years. The aim of this section is to provide the reader with a description and better understanding of Big Data and data monetization to avoid confusion and erroneous interpretations of the term.

Nowadays, several definitions of Big Data circulate in academia and the corporate world. Still, there is a large degree of consensus among scholars and practitioners around the definition provided by Doug Laney in 2001. Doug Laney came up with the characterization known as the 3 V's to describe the concept of Big Data and the entailing challenges that are specific to Big Data namely volume, velocity and variety. (Zhang & Chen, 2014)

The first and most known of the "V's" is *volume*. As the word itself indicates, Big Data alludes to big volumes of data. Hence, volume characterizes the size of the data. McAfee & Brynjolfsson (2012, p. 5) argues that "declining costs of all the elements of computing – storage, memory, processing, bandwidth, and so on" made data storage more economical allowing organizations and individuals to store larger data volumes than ever before. In like manner, Jagadish (2015, p. 52) emphasizes that data collection and storage became cheap "due to ubiquitous digitization, business process automation, the web, and sensor networks, (…) [and] falling media prices" As a consequence, many firms transform from a "data poor" to a "data rich" organization. (Jagadish, 2015)

In fact, the International Data Corporation (IDC) reported that the data and information volume created worldwide grew from 0.1 zettabytes in 2005 to 12 zettabyte in 2012, an increase by over two orders of magnitude (1 ZB equals 1 billion TB). Furthermore, IDC predicts that the global data volume will hit 47 ZB by 2020 and will reach over 160 ZB just five years later. (IDC, 2018) According to Wamba, Akter, Edwards, Chopin, & Gnanzou (2015), this rapid increase is also driven by the increasing adoption of mobile devices, Internet of Things, and social media platforms, such as Twitter and Facebook. Similarly, Habib ur Rehman, Chang, Batool, & Ying Wah (2016) assert that leading organizations increasingly collect massive amounts of internal and external data. However, these massive data sets can not be processed and stored in the traditional way, therefore, "… [qualifying], in terms of Volume, as Big Data" as argued by Jagadish (2015).

The second "V" is known as *velocity*. Nowadays, data is collected and processed in an ever faster manner. In 2018, every minute, over 4 million videos were viewed on YouTube, over $860,000 were spent online and over 260.000 hours were watched on Netflix (Lewis, 2018). Velocity refers to the rate at which data flows in and out of Big Data systems and the speed at which it is stored, processed and analyzed (Emmanuel & Stanier, 2016). According to McAfee & Brynjolfsson (2012), for many organizations, the speed of data creation is more relevant than volume since "real-time or nearly real-time information makes it possible for a company to be much more agile than its competitors." Yet, with velocity comes the Big Data challenge of being able to manage the high inflow rate of data (Sivarajah, Kamal, Irani, & Weerakkody, 2017).

The third "V" is *variety*. GPS signals, photos, videos, audio files, social media, mobile, log-file and web data are just an excerpt from a large variety of external and internal data sources. These data sources are collected in different formats, semantics and may be structured, semi-structured or unstructured. Scholars widely agree that variety is a key characteristic of Big Data and a challenge for many organizations given that traditional structured databases are incompatible of storing and processing different data formats (McAfee & Brynjolfsson, 2012). In like manner, advocates, such as Jagadish (2015) claim that this characteristic of Big Data is neglected in both academia and industry although it challenges organizations more than volume and velocity.

Doug Laney understood that Big Data architectures and processing are scalable and support the processing requirements of data high in volume, velocity, and variety. Since Laney's definition of Big Data, other characteristics have emerged over time. While scholars and practitioners discuss several extensions of the 3 V's, the following two have the highest consensus in both academia and industry. (Emmanuel & Stanier, 2016)

IBM declared *veracity* as the fourth "V" stressing the unreliability in some data sources. IBM supports this by stating that one out of three business leaders does not trust the information used, while 27% of respondents were unsure of the degree of inaccuracy of their data. Hence, IBM assumes that the costs of over $3 trillion per year are imposed on the economy due to poor data quality. (IBM, 2018) IBM's argumentation also receives support from scholars, such as White (2012, p. 211) who emphasizes "if data is not of sufficient quality by the time it has been integrated with other data and information, a false correlation could result in the organization making an incorrect analysis of a business opportunity."

Later, Oracle joined the debate and introduced a fifth "V" – *value*. Oracle argues that Big Data is often characterized as having "low-value density" in relation to its large volume. However, analyzing such large volumes allows for high-value creation. (Haider & Gandomi, 2015) This requires organizations to invest in sophisticated, large-scale data processing and analytics which, in the end, needs to pay off and provide a positive return on investment for organizations (Chiang, Grover, Liang, & Zhang, 2018). This leaves business and information technology leaders with the key challenge namely determining the value/pay off from data prior to analysis.

Furthermore, as asserted by Sivarajah, Kamal, Irani, & Weerakkody (2017), “old ways of data modeling no longer apply due to the need for unprecedented storage resources/capacity and computing power and efficiency” leaving academia and industry with the challenge to develop new methods to allow for more efficient and productive data processing in order to extract high value from Big Data. Consequently, throughout literature, it is undisputed that Big Data analytic capabilities, such as proper Big Data systems, analytical tools, and human talent as well as know-how, play a crucial role in creating value from Big Data.

Value can be measured and applied in different areas. For instance, it can improve business performance by increasing the efficiency of business processes and information flows. Additionally, generated insights can foster innovation in product and service development or improve customer loyalty and retention by leveraging, for instance, customer reviews and market data. Moreover, Big Data can generate value due to insights improving managerial decision-making which allows for e.g. better and faster responses to environmental and market changes. How to assess and derive quantifiable economic benefits, namely value, from Big Data is also known as *“data monetization”* in academic literature and debates. (Grover, Chiang, Liang, & Zhang, 2018)

It can be concluded that the term Big Data covers a wide range of data, applications, as well as technologies and should be seen in the context of its environment. Therefore, Emmanuel & Stanier (2016) claim that defining Big Data is challenging. Yet, the understanding of Big Data based on Laney's definition in combination with the extension of veracity and value is widely accepted among scholars and industry experts, thus, providing a solid foundation for the purpose of this paper. Furthermore, it is indisputable that capitalizing on Big Data bears the above-mentioned challenges and requires Big Data initiatives to be thoroughly thought through.

3. Methodology

In this paper, I applied an in-depth systematic literature review methodology. The objective of this paper is to provide the reader with a summary and analysis of the latest theories on Big Data monetization strategies based on the latest academic research articles and expert reports. According to Wamba, Akter, Edwards, Chopin, & Gnanzou (2015), relevant Big Data journal articles started to be published in 2011. Hence, I considered papers available since 2011 and up to December 2018 to provide insights of most recent date. Additionally, I seek to develop an overall understanding of Big Data for the reader.

First, I identified relevant English academic papers by performing an extensive keyword search through the "NOVA Discovery Meta-Search" tool. The "NOVA Discovery Meta-Search" tool contains, among others, full-text academic articles, eBooks, and books from UNL libraries. The comprehensive search contained the keyword "Big Data". The following subsequent keywords were used to further filter available academic papers: "value", "data monetization", and "data business model". The findings are drawn from publications in internationally renowned journals, such as the International Journal of Information Management, Journal of Strategic Information Systems, and the Journal of Management Information Systems. Given that journal quality is an indicator for research quality (Judge, Cable, Colbert, & Rynes, 2007), only high-quality peer-reviewed academic papers, have been used to provide the reader with a theoretical background.

Moreover, Hartmann, Zaki, Feldmann, & Neely (2016) found that contributions from academics and scholars mainly focus on the technical aspects, while most literature on the path to value creation through Big Data is written by consultancies and IT firms. Therefore, to expand the scope of this paper and include active debates on Big Data monetization, I conducted a comprehensive search through online sources, such as YouTube and Google, in order to generate new insights from recent Big Data conferences and Information Technology expert reports. The search selection was performed with the above-mentioned keywords. As a result, this paper provides a balanced perspective from academia and the practical approaches from industry leaders on data monetization.

4. Data monetization

Today, Big Data discussions in academia are still scarce, however, no longer solely subject to volume, variety, and velocity but more to the value of data. Yet, the path towards strategic data monetization remains unexplored and unclear for many organizations. Therefore, this section presents and discusses the relevant findings of this in-depth systematic literature review based on the latest academic research papers as well as expert reports dealing with the key question of *how* organizations can monetize Big Data. However, before this question is addressed, the discussion starts by providing the reader with a crucial understanding of the different strategic business fields where data can be monetized and add value to the organization, respectively.

Later, the findings on *how* organizations can strategically approach data monetization are discussed in section 4.2.

4.1. Potential of data monetization

To understand the potential fields of business application of data monetization, we draw on the findings from Wamba, Akter, Edwards, Chopin, & Gnanzou (2015). They performed a comprehensive literature review of journal articles about value creation from Big Data. Their widely-read findings from 62 relevant academic research articles conclude that scholars identify value creation potential in the following five business dimensions in the order of importance:

(1) *"Replacing/supporting human decision-making with automated algorithms"* which attributes to enhanced decision-making processes by reducing risks and extracting key insights. Such automation algorithms allow analyzing real-time data, thus, increasing situational awareness. For instance, Procter & Gamble is an American multinational consumer goods company that introduced algorithmic intelligence "Business Sphere" conference rooms for its leadership teams that feature a "visually immersive data environment that transforms decision-making at P&G by harnessing real-time business information from around the globe." (Summers, 2018)

(2) *"Enabling experimentation to discover needs, expose variability, and improve performance"* by collecting and analyzing data in real-time from controlled experiments to generate insights more quickly. To name an example, the American media-service provider, Netflix, continuously designs controlled experiments to

understand and test the viewer's perception towards new customer-centric media services and products (Ramirez, Frankwick, & Xu, 2016). In detail, Netflix analyzed 30 million plays, 4 million ratings, and 3 million searches to identify that viewers of the original House of Cards TV show predominately watched movies directed by David Fincher and starred by Kevin Spacey. Netflix discovered the need and experimented with an improved version of the original TV show. As a result, Netflix introduced the revised House of Cards show featuring Fincher as well as Spacey which became one of the most successful TV shows of Netflix.

(3) *"Innovating new business models, products, and services"* can be facilitated more effectively and productively when examining customer experiences and identifying product and service issues as well as desired features. Such insights not only improve next-generation product development but also enable companies to create new products and services as well as new business models. For instance, the Spanish fast-fashion retailer Zara performs real-time analysis of popular fashion trends from external data as well as customer spending from internal data to provide its customers with the latest and most trendy fashion (Sorescu, 2017).

(4) *"Segmenting populations to customize actions"* is utilized by companies to create highly specific segmentations and to personalize products and services to meet individual customer needs precisely. The hospitality industry, in which personalization and experience is a key differentiator, aims to provide the highest level of comfort and personalization for their guests through Big Data analytics. For example, Hilton is actively personalizing guest experiences by "…[setting] it [the room temperature] for you before you get there. Equally, and very important, the entertainment system in our

hotels, (we will be) personalizing that so that if you watch your preferred channels, CNBC (for example), those will be on your TV when you walk into the room." stated Hilton Chief Marketing Officer Geraldine Calpin (Handley, 2018).

(5) *"Creating transparency"* reduces search and processing procedures, accelerates time to market, and enhances quality standards. For instance, as opined by Grover, Chiang, Liang, & Zhang (2018), transparency enables consistency in viewing data across the organization and facilitates the visibility of internal business processes through, for example, real-time dashboards. Another example is provided by Günther, Rezazade, Huysman, & Feldberg (2017) who argue that a growing number of organizations and governments opt to make their data publicly available in order to foster innovation. The Bank of England is an example of embracing transparency given its approach to disclose previously confidential datasets to the public in order to crowdsource solutions for its policy issues.

For the present purpose of this paper, it is sufficient to provide the reader with a universal overview of the potential business applications of data monetization. Those five universal dimensions provide a profound overview of the potential strategic business applications of data monetization. Wang (2012) further investigated the dimension of "Innovating new business models, products, and services" and found that the following three approaches are most common among industry leaders. First, creating new customer experiences through strong differentiation. This aims at helping businesses to not only develop new product and service offerings but also to satisfy customers while providing contextual relevance. Second, brokering raw information will evolve into "sub-sub specialized" information streams that provide even deeper insights and analysis for buyers of such raw data. For example, Facebook's facial recognition algorithms can identify the top 10% of users that have gained most weight over the last year. Facebook can sell this information to e.g. Weight Watchers to place an advertisement on these users' news feed. (Marr, 2016) Third, delivery networks aggregate data and provide a marketplace for organizations to exchange and to trade relevant data. The development of delivery networks, however, requires vast capital and technology resources.

While the systematic literature review of this research paper has not recognized new dimensions from academics and scholars since the publication of Wamba, Akter, Edwards, Chopin, & Gnanzou (2015), an interesting contribution was identified from McKinsey & Company, a global American management consulting firm, in 2017. The above-mentioned dimensions focus primarily on data monetization potential of a single company. McKinsey & Company investigated the value realization effect among two or more entities. As stated by Gryseels (2017), McKinsey & Company introduced the "1+1=3" model which argues that organizations monetize data in highly sophisticated use cases through collaboration. To illustrate, insurance providers struggle to tailor insurance policies to individual customer needs despite the increasing demand for personalized policies. Through partnerships with telecommunication companies, those insurance providers can utilize telco data to analyze, for example, the driving behavior of customers based on GPS data to provide individualized car policies. In fact, Uber, an American peer-to-peer ride-sharing provider, announced "the first collaboration of its kind" (GovLab, 2018, p. 1) with the City of Boston. The idea behind this partnership is that Uber offers anonymized trip data of its users to the City of Boston in order to support the municipal with insights on traffic congestions, public transportation issues, urban growth, and CO2 emissions. With the help of these insights, city planners make more informed decisions and address key issues more effectively. (GovLab, 2018)

In like manner, when two or more companies are involved in data collaboration, McKinsey & Company talks about “data ecosystems”. Their exhaustive industry assessment identified 12 data ecosystems, such as education, B2C marketplace, and health, in which data collaboration among organizations allows for exponential value creation. In fact, they estimate the global value of such ecosystems at USD 60 trillion. Gryseels (2017) argues that companies in such ecosystems lack the capability to create such value individually given that such highly valuable insights cannot be generated without the data of other firms. In fact, Ping An, a Chinese banking and insurance company, operates in five data ecosystems today. For example, their “Good Doctor” platform created an ecosystem in the health industry attracting over 75 million users (O'Dwyer, 2018). As a result, Ping An has more health data of Chinese than any other public or private company in China. This abundance of data allows Ping An not only to advice their customers on treatment specific doctors and hospitals but also on their insurance policies tailored to individual customer needs. (Gryseels, 2017)

All in all, based on latest contributions from academics and industry experts, the discussion of this section provides the reader with a clear direction of where the different fields of data monetization can contribute to organization’s business model. In addition to the afore-mentioned and widely-discussed five dimensions, McKinsey & Company’s findings on data monetization through ecosystems were added to introduce the latest practical approaches from industry/Big Data leaders.

4.2. The path to data monetization

The path to monetizing data remains a challenging task for many organizations. Not only the technical complexity and the most often needed organizational changes, but also high initial investment costs without certain outcome make this endeavor highly risky. As a result, the number of organizations that have successfully walked the path to monetizing data remains scarce. (Chen, Schütz, Matthes, & Kazman, 2017) While the previous section developed a basic understanding of the different business fields in which data monetization can be realized, this section, now, aims to answer the remaining question and shed light on the limited understanding of *how* organizations can unlock the value of data and monetize it.

Few but important studies have been investigating the paradigm of data monetization (e.g. Chen, Schütz, Matthes, & Kazman (2017); Günther, Rezazade, Huysman, & Feldberg (2017); Grover, Chiang, Liang, & Zhang (2018)). Lim, et al. (2018) provide organizations with a nine-factor framework to describe, analyze and design the value creation process in the information-intensive service industry. Although empirical studies are still required to assess the relevance of this framework for other industries, it is assumed that four out of the nine factors are universal and deliver a basic understanding of the initial process from data to monetization for the reader. Lim, et al. (2018) argue that the value creation process starts with the (1) data source. Organizations must understand where they need to collect the desired data, which in turn requires organizations to have a predefined strategic Big Data goal. The next step is (2) data collection, which suggests that organizations also need to determine the appropriate methods to collect data. Furthermore, organizations must be clear about the content and what kind of (3) data is collected. The next factor (4) data analysis refers to the algorithms and tools used to extract the value from data. In terms of data analysis, Grover, Chiang, Liang, & Zhang (2018) published a widely-read conceptual framework for value creation. The authors stress that it is fundamental for organizations to build a Big Data analytics infrastructure by investing primarily in human talent and Big Data assets. This infrastructure should have "strong capabilities of integrating, managing, sharing, and analyzing Big Data in diverse formats…" (Grover, Chiang, Liang, & Zhang, 2018, p. 399) In like manner, Najjar & Kettinger (2013) investigated the path to data monetization of a major U.S. drug retailer. They found that this American drug retailer built its technical data infrastructure first. The strategic intent behind this approach was to look at data as an asset and trying to collect as much data as possible as early as possible which was later monetized by selling it to supply-chain partners. In addition, Najjar & Kettinger (2013) describe two alternative pathways to data monetization based on technical, such as infrastructure and hardware, and analytical capabilities of the organization: (1) Building technical and analytical capabilities internally at the same time. This, however, bears the highest risks and requires vast initial investments. (2) Building analytical capabilities first. This approach focuses on developing human analytics talent first, and later in the process the required technical capabilities.

Equally important to technical capabilities and process changes are organizational changes. Organizations need to develop stark leadership and corporate culture embracing data-driven processes given that data monetization involves several stakeholders and departments within the organization. Additionally, in the end, human beings still need to accept, decide and implement insights into strategic and operational decisions to generate value. Therefore, it is essential that the organizational routines, structures, and decision-making processes are aligned for data monetization. (Sharma, Mithas, & Kankanhalli, 2014) In fact, an increasing number of companies change their organizational structure and establish a Chief Data Officer (CDO) role to manage data (monetization) processes. CDOs evolve as key leaders in data-driven organizations as they align Big Data initiatives with the overall business strategy and oversee data governance and compliance. Furthermore, CDOs are executive-ranked, thus, able to foster direct communication with the CEO while ensuring top management support and close executive collaboration for data initiatives. (Lee, Madnick, Wang, Wang, & Zhang, 2014)

In another study, Günther, Rezazade, Huysman, & Feldberg (2017) performed a comprehensive literature review of 67 academic papers and argue that many organizations collect data without a pre-defined strategic goal which results in an inductive (bottom-up) data monetization approach. Scholars, however, strongly advice against an inductive approach since such an approach is thought to be expansive and resource heavy. Instead, they propose a deductive (top-down) approach which starts with a general hypothesis and pre-defined goal. Data is then used to test the hypothesis. Conducting a deductive approach and having a pre-defined goal and hypothesis is "essential for maximizing the likelihood of value realization." (Tamm, Seddon, & Shanks, 2013, p. 12). Alphabet Inc., the parent company of Google, for example, developed a list of 35 most important strategic business questions evolving around major business challenges. Any management decision, including Big Data initiative, has to have a link to these high-level unanswered strategic questions in order to get the board approval (Marr, 2016). A top-down approach can also reduce the costs of Big Data analysis given that the utilization costs of cloud services for data analysis are increasing (Habib ur Rehman, Chang, Batool, & Ying Wah, 2016). Therefore, Habib ur Rehman, Chang, Batool, & Ying Wah (2016) argue that Big Data reduction measures are also key to value creation. They stress that organizations need to perform initial data reduction operations prior to analysis to reduce cloud utilization costs, and to increase data privacy for users while securing data sharing to build trust between users and organizations.

A deductive approach was applied by Lufthansa, the largest German airline company, which, according to Chen, Schütz, Matthes, & Kazman (2017, p. 20), is "the Amazon of the airline industry" due to its highly successful transformation to a data-driven business model. Prior to designing and building a Big Data infrastructure, Lufthansa performed a "Top-Down Value Discovery Process" which consisted of three phases: (1) innovation process, (2) use case development and (3) strategic development planning. This approach generated hundreds of use case which underwent a selection process involving prioritization, cost-benefit and feasibility assessments as well as prototyping. As a result, they identified four key use cases, which enabled Lufthansa to generate value through Big Data. Another key factor that contributed to the success of the Big Data initiative was the direct involvement and strong support from top management. (Chen, Schütz, Matthes, & Kazman, 2017)

Lufthansa's approach to data monetization is comparable to McKinsey & Company's proposed path to data monetization. Senior Partner, Gryseels (2017), developed a four-step journey to data monetization. The first step requires the company to start an in-depth process to identify and prioritize the use case(s) that generate the highest value and low resource and capability requirements for implementation. The second step involves designing effective data governance with clear responsibilities and building a modern data architecture which is supposed to consist of only one data lake where data is aggregated and accessible to the entire organization. The third step refers to building the (human) capabilities needed to monetize data. In detail, organizations need capabilities in business, such as business leaders who can frame data problems and understand the technical and business side; in technology, such as data engineers who can integrate and manage Big Data; and in analytics, such as data scientists who can extract insights. In fact, McKinsey & Company found that for each data scientist, on average, five data engineers are needed. Lastly, (Gryseels, 2017) stresses the importance of developing a corporate culture that embraces agile working approaches and a data-driven environment as well as top-management support.

5. Concluding remarks

The objective of this paper was to shed light on the limited understanding of how organizations can unlock the value of data and monetize it by drawing on the findings from an in-depth systematic literature review of academic papers on Big Data. On the basis of this study, it is obvious that the current literature on value creation from Big Data is limited. Scholars and academics have defined various paths to data monetization but have primarily focused on the technical mechanisms, processes and analytical tools of Big Data. Still, based on the synthesis of relevant academic papers and practical inputs from business leaders, it can be concluded that organizations are well adviced to choose the path to data monetization which embraces a deductive approach. More and more organizations and business leaders argue for having, at first, a general hypothesis and pre-defined goal (Marr, 2016). Afterward, use cases are identified and assessed based on costs and feasibility and later tested with data. This process ensures that there is an actual business value behind the use case and data prior to investing in Big Data infrastructures and analytical capabilities. While the approach to data monetization may vary from industry to industry and organization to organization, the importance of fostering organizational changes towards a company-wide culture that embraces data-driven processes and encourages as well as rewards data-driven decisions, is a key factor for any organization in order to ensure the success of data monetization initiatives.

5.1. Limitations and future research

In this paper, the nature of the applied research methodology leads to limitations that readers and future researchers, as well as academics, should be aware of. The first limitation of this study stems from the literature search through the "NOVA Discovery Meta-Search" tool, which was limited to English speaking publications in a restricted number of databases. Second, the selection and filtering for relevant academic papers was based on specific keywords in the title, abstract or body. Therefore, this paper addresses explicitly the field of Big Data and cannot guarantee exhaustiveness as some relevant research articles might use different keywords and were not identified in this review. Lastly, since academic papers and empirical studies on data monetization are scare, there is little literature to build on. Hence, this paper relies on observation and insights from successful firm practices which have not been empirically tested.

Therefore, future research is needed to further investigate the pathway to data monetization. While businesses can draw on various academic contributions on technical aspects of Big Data and data monetization, there has been little research on how these organizations can translate data into strategic business value. Scholars and researchers should address this gap in the future.

6. References

Accenture. (2018). *Accenture Technology Vision 2018.* Accenture.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence And Analytics: From Big Data To Big Impact. *MIS Quarterly*, 1165-1188.

Chen, H.-M., Schütz, R., Matthes, F., & Kazman, R. (2017). How Lufthansa Capitalized on Big Data for Business Model Renovation. *MIS Quarterly Executive*.

Chiang, R., Grover, V., Liang, T.-P., & Zhang, D. (2018). Special Issue: Strategic Value of Big Data and Business Analytics. *Journal of Management Information Systems*, 383-387.

Emmanuel, I., & Stanier, C. (2016). Defining Big Data. *ACM International Conference Proceeding Series* (pp. 1-6). ACM.

Gartner Inc. (2018, November 13). Retrieved from Gartner Newsroom: https://www.gartner.com/newsroom/id/3130017

Google Trends Explore. (2018, November 12). Retrieved from Google Trends: https://trends.google.com/trends/explore?date=all&geo=US&q=big%20data

GovLab. (2018, November 30). Retrieved from Uber-City of Boston Partnership: http://datacollaboratives.org/cases/uber-city-of-boston-partnership.html

Grable, J. E., & Lyons, A. C. (2018). An Introduction to Big Data. *Journal of Financial Services Professionals*, 17-20.

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

Gryseels, D. M. (Director). (2017). *Big Data Monetization* [Motion Picture].

Günther, W., Rezazade, M., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems*, 191-209.

Habib ur Rehman, M., Chang, V., Batool, A., & Ying Wah, T. (2016). Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management*, 917-928.

Haider, M., & Gandomi, A. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 137-144.

Handley, L. (2018, November 26). *CNBC*. Retrieved from Hilton hotels chief marketer : We want to personalize bedrooms using digital technology: https://www.cnbc.com/2017/09/15/hilton-hotels-chief-marketer-we-want-to-personalize-bedrooms-using-digital-technology.html

Hartmann, P., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing value from big data - a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations & Production Management*, 1382-1406.

IBM. (2018, November 14). *IBM Big Data Hub*. Retrieved from The FOUR V's of Big Data: https://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg

IDC. (2018). *Internet Trends 2018.* Kleiner Perkins Claufield & Byers.

Jagadish, H. V. (2015). Big Data and Science: Myths and Reality. *Big Data Research*, 49-52.

Judge, T., Cable, D., Colbert, A., & Rynes, S. (2007). What causes a management article to be cited—article, author, or journal? *Academy of Management Journal, 50*(3), 491-506.

Lee, Y., Madnick, S., Wang, R., Wang, F., & Zhang, H. (2014). A Cubic Framework for the Chief Data Officer: Succeeding in a World of Big Data. *MIS Quarterly Executive*.

Lewis, L. (2018, November 14). *Cumulus Media*. Retrieved from Visual Capitalist: http://www.visualcapitalist.com/internet-minute-2018/

Lim, C., Kim, K.-H., Kim, M.-J., Heo, J.-Y., Kim, K.-J., & Maglio, P. (2018). From data to value: A nine-factor framework for data-based value creation in information-intensive services. *International Journal of Information Management* , 121-135.

Marr, B. (2016). Big Data in Practice - Predict 2016. *Predict.* Dublin: Predict - Europe's Leading Data Conference.

Mazzei, M. J., & Noble, D. (2017). Big data dreams: A framework for corporate strategy. *Business Horizons*, 405-414.

McAfee, A., & Brynjolfsson, E. (2012). *Big Data: The Management Revolution.* Harvard Business Review.

Mikalef, P., Pappas, I., Krogstie, J., & Giannakos, M. (2016). Big Data and Strategy: a Research Framework. *Tenth Mediterranean Conference on Information Systems (MCIS).* Paphos: ResearchGate.

Najjar, M., & Kettinger, W. (2013). Data Monetization: Lessons from a Retailer Journey. *MIS Quarterly Executive*.

O'Dwyer, M. (2018, December 1). *The Digital Insurer*. Retrieved from China In-Depth: Digital Insurance Ecosystems: https://www.the-digital-insurer.com/china-in-depth-ecosystems-in-china/

Partenrs, N. (2018). *Big Data Executive Survey 2018.* NewVantage Partners.

Ramirez, E., Frankwick, G., & Xu, S. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research*, 1562-1566.

Schmidt, E. (2018, November 15). *Business Insider Nordic*. Retrieved from Business Insider: https://nordic.businessinsider.com/google-eric-schmidt-countries-will-fight-over-big-data-alphabet-cloud-2017-3/

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organizations. *European Journal of Information Systems*, 433-441.

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 263-286.

Sorescu, A. (2017). Data-Driven Business Model Innovation. *Journal of Product Innovation Management*, 691-696.

Summers, N. (2018, November 26). *Procter & Gamble latest innovations.* Retrieved from Procter & Gamble: https://www.pg.com/en_US/downloads/innovation/factsheet_BusinessSphere.pdf

Tamm, T., Seddon, P., & Shanks, G. (2013). Pathways to Value from Business Analytics. *Thirty-Fourth International Conference on Information Systems.* Milan.

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How ‘big data’ can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal Production Economics*, 234 - 246.

Wang, R. (2012). What a Big-Data Business Model Looks Like. *Harvard Business Review*.

White, M. (2012). Digital Workplaces: vision and reality. *Business Information Review*, 205-214.

Zhang, C.-Y., & Chen, C. P. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Science*, 314 - 347