MASTER

Why SMEs lag behind in big data
an explorative study of Dutch consultancy firms

Bokdam, A.

Award date:
2019

Why SMEs lag behind in big data

An explorative study of Dutch consultancy firms

By Abel Bokdam
Student number 1020148

Written in partial fulfillment of the requirements for the degree of

Master of Science in Innovation Sciences

At the Department of Industrial Engineering & Innovation Sciences
Eindhoven University of Technology

Supervised by
Dr. Bert Sadowski
Dr. Carolina Castaldi
Prof. Dr. Floor Alkemade

Title:
Why SMEs lag behind in big data

Subtitle:
An explorative study of Dutch consultancy firms

Keywords:
Big data, SMEs, organizational resources, information technology adoption

Author:
Abel Bokdam

Supervisors:
Dr. Bert Sadowski (Faculty of Industrial Engineering and Innovation Sciences. TU/e)
Dr. Carolina Castaldi (Faculty of Industrial Engineering and Innovation Sciences. TU/e)
Prof. Dr. Floor Alkemade (Faculty of Industrial Engineering and Innovation Sciences. TU/e)

Place, date:
Eindhoven, 21st of March 2019

Preface

This preface represents the final step of a long journey. When I started with the process of writing this thesis, I did not foresee the complications and obstacles I would face. I heard plenty of horror stories, but I never imagined that any of it could apply to me. Maybe I was naïve. Maybe I was still with my mind in New Zealand. Nevertheless, the process that was started in February 2018 after my return from Wellington is drawing to a close.

After such a long period, I am excited to be able to move on. This would not be possible however, with the help and support of various people. First, I would like to thank Bert for helping me (re)write and structure significant parts of the thesis. This forced me to think about and better scope my focus, which benefitted the thesis. Second, I would like to thank Carolina for her advice and support. You have always managed to find time for a meeting, be it on Skype or Utrecht Centraal

I would also like to thank the people at Technopolis Group Amsterdam. While my internship did not yield the data I was hoping for, it was instrumental for getting familiar with the topic. It was amazing to get to know so many knowledgeable people over more than a few coffees. Everyone was willing to think along with my thesis, but I also learned about a lot of other subjects. The fact that I could tag along and work on other projects made me feel part of the firm, especially as I was able to join the yearly Technodays in Brighton.

Finally, I would like to thank my family, girlfriend and friends for the support they gave during the long process. They have shown interest, listened to my complaints, reassured me everything would be fine, helped me get further and overall made it possible that this is the closing chapter for my time as a student. I am glad that I have people I can count on and plan to be there for them as well.

Abel Bokdam
March 2019

Summary

Data is increasingly seen as a source of competitive advantage and evidence based decision-making. With the decreasing costs of data storage and the datafication of life, more data is generated than ever. It is thought that these large datasets contain information and correlations that would otherwise remain concealed. However, processing this data remains difficult. This has given rise to the term 'big data' and new technologies to process it. Currently, large firms are taking the lead in unlocking the value of big data. Small and Medium-sized Enterprises (SMEs) are lagging behind, even though these firms could also strengthen their business by utilizing big data.

This thesis studies why SMEs are lagging behind by investigating big data use in the Netherlands. First, an overview of the concept of big data will be given. Then, a conceptual framework is constructed to see what is already known about the factors influencing big data use in SMEs. These factors form the basis of the interview protocol. Qualitative interviews are held to analyze what the perceived barriers and enablers were in the adoption and use of big data. Based on this research some theoretical and managerial implications are discussed.

Theory

Big data consists of information assets that are either high in volume, variety or velocity. Big data that is high in volume means that the data files are large in size and require extra processing power. The variety of data types is the result of the datafication process. Audio, video and unstructured text can also be analyzed, but do require new techniques. Finally, the velocity of big data is the speed in which it is gathered. Real-time data generation blurs the line between collection and analysis.

The resource-based view is used to understand why big data is framed as a driver for innovation and creator of a competitive advantage. It is conceptualized as a unique resource, which is difficult to replicate, imitate or substitute. In addition, big data could foster the development of IT capabilities and organizational knowledge, which enhances the firm's ability to leverage resources. This is what distinguishes big data from traditional IT investment. As a result, it can play a crucial role in creating and sustaining a competitive advantage for firms. To understand what influences the decision to adopt and use big data, a conceptual framework was constructed with factors from a literature review.

Methods

The research approach in this thesis is qualitative. Data is collected through interviews with representatives from consultancy SMEs in The Netherlands. These firms have business models based on knowledge, which big data could supplement. SMEs were approached via telephone, e-mail or LinkedIn and selected based on information on their website. Some firms

actively marketed themselves as big data users, while others provided big data services to third parties.

Representatives from these firms provide a unique perspective on how their firms went to the adoption process and what they perceived as the main barriers and enablers. The representatives held a management position or were involved in data science and engineering. This combined insights from people with knowledge of big data and executive decision-makers.

The 14 interviews were guided with the use of an interview protocol, which in turn was based on the conceptual framework. This scoped the research, but left room in the semi-structured interviews to deviate from certain questions. The quality of the research was safeguarded with a variety of measures, including triangulation of data sources, peer debriefing, member-checking and attending an expert workshop.

Results and Conclusion

There was no agreed upon exact definition of big data. The respondents used large databases or big data that differed in size and goals. This is partly due to fuzziness of the term, but nine respondents did fit the definition of this research. However, some evidence suggests that merely using the 3V model is not sufficient to capture the essence and complexity of big data.

In the successful adoption and use of big data, several factors emerged as barriers and enablers. The most often cited factor was human resources. This was believed by most to be the most important barrier to successful use of big data. As big data is complex, it requires expertise and specialized skills that are difficult to acquire. As a result, SMEs have trouble finding and retaining the right people to adopt big data technologies. The lack of resources of an organization can therefore be seen as a major barrier.

An important enabler was the flexibility of the organization in adjusting to change and incorporating new innovations like big data in the organization. An informal culture that is flexible, curious and where experimentation is valued were named important in being outward facing and successfully identifying opportunities like big data. This refers to the dynamic capabilities of a firm and the ability for change. Other factors of influence were concerns regarding the security and privacy of data, rules and regulations and the type of firm that is adopting big data. Especially firms that are B2C and have access to large volumes of data could benefit from big data technologies.

Implications

Reflections on the findings indicate that big data might not be as unique as claimed. The barriers and enablers that emerged from this research do not differ significantly from those in more general IT studies. Furthermore, there was inconclusive evidence that big data was used as a method for developing IT capabilities. However, this does not mean that big data does not create any value. The firms currently utilized it mostly for enhanced decision-making and as a resource or product. This has implications for both managers and policy makers.

For managers it is important to stimulate employees in experimenting with new technologies. These people form the eyes and ears of a firm and can help in identifying problems and opportunities. The lack of skilled people for hire means that the assets of the firm should be used to support skill-development in-house. While the short-term benefits of big data use might not be apparent, it could become a sustainable competitive advantage. Distinguishing yourself from the competition or optimizing processes are some of the traditional advantages of big data.

For policy makers it is essential to stimulate students to choose educational programs that pay attention to big data and digital skills. In the long term, it is essential to educate, upskill and reskill people to close the skills gap. Furthermore, the number of graduates should not only be increased, but measures need to be taken to make data science attractive for a more diverse group of people. Big data gives rise to new ethics concerns, which are viewed differently by engineers and social scientists.

Contents

List of abbreviations

B2B	Business to Business
B2C	Business to Consumer
BDA	Big Data Analytics
BDAaaS	Big Data Analytics as a Service
CRM	Customer Relationship Management
DC	Dynamic Capabilities
DOI	Diffusion of Innovations
EC	Environmental Characteristics
GB	Gigabyte
GDPR	General Data Protection Regulation
IC	Innovation Characteristics
INT	Interviewee
IoT	Internet of Things
IS	Information System
IT	Information Technology
MVP	Minimum Viable Product
OC	Organization Characteristics
RAM	Random Access Memory
RBT	Resource Based Theory
RBV	Resource Based View
RQ	Research Question
SME	Small and Medium Enterprise
TB	Terabyte
TOE	Technology-Organization-Environment
VRIN	Value, Rareness, Imitability, Non-Substitutability

List of figures

List of tables

Chapter 1 – Introduction

Big data is presented as the next driver for innovation, competition and productivity (Manyika et al., 2011; Porter & Heppelmann, 2014). A wide range of industries could benefit from big data replacing human decision-making, generating new business models, customization of products and services, increasing segment understanding, uncovering needs and much more (Cavanillas, Curry, & Wahlster, 2016; Manyika et al., 2011; Raguseo, 2018). The promise of all these possibilities reminds of the emergence of Information Technology (IT) in the late 20th century (Oliner & Sichel, 2000). Similar to big data now, IT would revolutionize productivity and drive innovation (Brynjolfsson, 1993).

Studies on the link between IT investment and firm performance found mixed results, explaining that IT acquisition alone is not enough for an increase in firm performance (Tippins & Sohi, 2003; Wu, Yeniyurt, Kim, & Cavusgil, 2006). This apparent productivity paradox has been studied extensively (Brynjolfsson & Hitt, 1998; Tippins & Sohi, 2003). An explanation is found in the framework of dynamic capabilities (Teece, Pisano, & Shuen, 1997). These are required to integrate, build and reconfigure organizational resources to exploit them for a competitive advantage (Teece et al., 1997).

However, the current interest in big data disregards the established academic relationship between IT and productivity, as big data is claimed to be different (Hammer, Kostroch, & Quirós, 2017). Compared to previous data analytics, big data is higher in the volume, velocity and variety of data available, collected and analyzed (McAfee & Brynjolfsson, 2012). This is facilitated by the increase in computing power combined with decreasing data storage costs, which have unlocked the potential of datafication (Hammer et al., 2017).

The focus is on the impact these developments have for the strategy and competitive advantage of firms (Mazzei & Noble, 2017; Porter & Heppelmann, 2015; Raguseo, 2018). Firms can differentiate themselves by transforming into pro-active and forward looking organizations based on big data (Wamba et al., 2017), but little is known about the issues firms face in leveraging their resources and capabilities to actually reach that stage.

1.1 Problem statement

The use of data by firms is nothing new, as data already impacts the European economy as a whole (Brynjolfsson & McElheran, 2016; IDC & The Lisbon Council, 2018b). Within countries of the European Union, The Netherlands was ranked highest in the global competitiveness index of 2016 and 2017 (Schwab, Sala-i-Martín, & World Economic Forum, 2017). Large sectors of the country's economy are based on services and IT (Centraal Bureau voor de Statistiek, 2018) and it has the highest percentage of firms using data (IDC & The Lisbon Council, 2018a).

In particular service industries like consultancy capitalize on information (Turner, 1982) and could benefit from big data (Christensen, Wang, & van Bever, 2013). Consultancy firms are knowledge intensive business services that depend on the skill and knowledge of its personnel and the information they have (Cesário, Fernandes, Jesus, & Barata, 2015; Obeidat, Al Suradi, Masa'deh, & Tarhini, 2016). For these firms, big data poses both a threat and an opportunity. If utilized it can enhance the firm's credibility and position. However, data is also seen as a possible disruptor of the consultancy business by unpacking the 'black box' in which recommendations are made, enabling clients to better understand and evaluate the quality of the services (Christensen et al., 2013).

When looking at the big data use among Dutch firms more closely, it appears that smaller firms are lagging behind (Centraal Bureau voor de Statistiek, 2017). This is worrisome, as small and medium enterprises (SMEs) form the backbone of any economy and are often associated with flexibility and innovation (Dosi, Moschella, Pugliese, & Tamagni, 2015; Gray, 2006). If big data is truly the driver for competitiveness in all industries (Manyika et al., 2011), then SMEs have to start utilizing big data more (Coleman et al., 2016).

In summary, most studies currently focus on the impact of big data on performance or competitiveness of large firms (Mazzei & Noble, 2017; McAfee & Brynjolfsson, 2012; Wamba et al., 2017) in industries with tangible outputs (Manyika et al., 2011; Porter & Heppelmann, 2014; Sen, Ozturk, & Vayvay, 2016). An understanding of how firms leverage resources and capabilities to create a competitive advantage is lacking, in particular for SMEs in consultancy-service industries. Investigating these issues is crucial in understanding why SMEs lag behind in the adoption of big data.

1.2 Research questions

This research aims to contribute in understanding why SMEs lag behind in leveraging big data for a competitive advantage. The contribution will be twofold. The first contribution will be in understanding of what sets big data apart to create a competitive advantage. The second and main objective is to understand what issues SMEs face in deploying their organizational resources and capabilities to adopt big data in The Netherlands. The result will be a scientific overview of factors that influence the decision and ability to use big data by SMEs. The particular contribution of this research is in the understanding of consultancy-service SMEs and whether big data is truly different from other IT. As big data might transform both companies and a wide range of industries (Porter & Heppelmann, 2014, 2015), understanding of this subject is essential for informed decision-making by both managers and policy makers. Based on this and the previous paragraphs, the main research question is formulated as:

What makes big data unique and how do organizational resources impact its adoption by Dutch small and medium consultancy SMEs?

In this thesis, the main research question is exploratory in nature and will be divided into four sub-questions. First, it is necessary to understand what big data is. This leads to the first sub-question, which is theoretical in nature:

Sub-question 1: *What is big data?*

The second question is about determining why big data is a unique resource and if it differs from traditional information technologies. This will add to the understanding of how big data could transform companies and how it could play a role in creating a competitive advantage according to the literature. This question is part theoretical and empirical in nature, to also get an indication of the current practices of big data. The question is formulated as:

Sub-question 2: *Why is big data a unique resource and how does it differ from traditional IT in Dutch consultancy SMEs?*

After the theoretical part of research question two, big data is framed in the context of organizational resources and firm performance. The next step is to construct a conceptual framework to study why SMEs are lagging behind in the use of big data. A literature review will be used to construct a conceptual framework and to identify potential factors that influence the adoption and use of big data. The third question is conceptual in nature:

Sub-question 3: *What are the potential challenges of big data adoption?*

The final sub-question will investigate what factors influence the adoption of big data by Dutch consultancy SMEs and is based on the conceptual framework from the third research question. A better understanding of the processes that could hinder or stimulate the successful use of data within Dutch consultancy-service firms is vital for answering the main research question. The fourth question is therefore:

Sub-question 4: *What factors influence the adoption of big data by Dutch consultancy SMEs?*

1.3 Relevance

Big data is an emerging field and still very much in development (Cavanillas et al., 2016). As a result of the hype surrounding big data, research-funding and research on the topic is rapidly increasing (Al-Qirim, Tarhini, & Rouibah, 2017; Davies, 2016; Riggins & Wamba, 2015). However, an extensive agenda for future research remains, especially empirical studies on big data (Loebbecke & Picot, 2015; Riggins & Wamba, 2015). At the same time, researchers point to the fact that a clear and widely agreed upon definition of big data is lacking (Cavanillas et al., 2016; Hammer et al., 2017). Without a working definition the concept of big data is fuzzy, fragmenting the academic literature and slowing the accumulation of knowledge.

This is apparent in the many articles heralding big data as the new source of competitive advantage (Gunasekaran et al., 2017; Hartmann, Zaki, Feldmann, & Neely, 2016; McAfee & Brynjolfsson, 2012; Raguseo, 2018; Wamba, 2017). Between these studies the definition of big data varies, making informed discussions and comparing key insights more difficult. For example, several articles claim that all firms, small or large, could benefit from big data (Coleman et al., 2016; Shah, Soriano, & Coutroubis, 2017). Without agreement on what big data actually is, it is difficult to value this claim.

The scientific relevance of this thesis is to contribute to an understanding of why SMEs lag behind in big data and which organizational resources play a role. This requires a discussion on what big data is to guide the research. In doing so, this thesis will characterize big data in terms of organizational resources, which contributes towards understanding the link between big data and competitive advantages. A second contribution is the construction of a conceptual framework to study factors of influence on big data use. This framework is then applied to the case of consultancy SMEs in The Netherlands for a better understanding of the reality of big data use.

This leads to the societal relevance of this thesis. The European Commission identifies the big data adoption lag by SMEs as the biggest issue for the future competitiveness of the European economy (European Commission, 2016b). It attributes the lag to the lack of skilled and knowledgeable people capable of working with big data in the European workforce (Cedefop, 2018; European Commission, 2016a). In turn, the skills gap will especially hit SMEs as they are unable to pay the premium for skilled people (Coleman et al., 2016).

In response, the European Commission initiated a project that is meant to design and test measures to close the digital skills gap for SMEs (Executive Agency for SMEs, 2017). This research is partly carried out within the overarching project as part of an internship at the Technopolis Group. However, the skills gap is only a partial answer to the lag in big data use by SMEs. More skilled people do not necessarily have to lead to more big data use in small firms. A more holistic approach is needed to understand what influences the adoption of big data technologies by SMEs.

It is necessary to deconstruct the hype currently surrounding big data by providing a clear definition. In addition, the findings of this thesis help understand the current situation of big data usage in Dutch consultancy SMEs, which could inform entrepreneurs, firm-management and policymakers.

1.4 Research Design

In this thesis I take a qualitative approach to answer the research questions. A qualitative approach is suited, because the focus is on how organizational resources could influence the effective deployment of big data. This is context dependent due to the diversity of firms and use-cases. As a result, tackling these issues with a qualitative approach enables me to dive deeper into specific cases while taking into consideration the unique context.

A deductive research design is chosen as the result of the broad concepts that are being studied (Creswell, 2014). A conceptual model is used as a foundation of and guide to the research. This narrows the scope of the research from the general to the more specific. This enables the research to build upon the existing body of knowledge on the resource-based view and dynamic capabilities.

Together the sub-questions provide an answer to the main research question *what makes big data unique and how do organizational resources impact its adoption by Dutch small and medium consultancy SMEs?* Starting with the first research question *what is big data?* I will answer this question with the use of the existing academic literature on big data to define the concept.

The second research question is w*hy is big data a unique resource and how does it differ from traditional IT in Dutch consultancy SMEs?* This question has two parts. First, by discussing the resource-based view and dynamic capabilities an answer can be formulated as to why big data is a unique resource. Resources have certain characteristics that determine how valuable they are for the firm in creating or sustaining a competitive advantage. However, the ability to exploit these resources is just as important. By discussing the characteristics of big data it is possible to ascertain why it is a unique organizational resource and how it could help in fostering capabilities. This frames it in the larger context of creating a competitive advantage.

The second part of this research question is necessary to ascertain whether big data is truly different compared to traditional IT. It provides an indication of how big data is currently applied, but also provides insight in the merit of the claims that big data is transformational. This is essential for the fourth research question, as these challenges might be different than regular IT projects. This part of the research question is empirical in nature.

The third research question is *what are the potential challenges of big data adoption?* This conceptual question performs a literature review to identify the factors that influence the adoption and use of big data. It forms the conceptual framework, which will be used in the empirical investigation of this thesis. The overview of challenges enables better scoping of this research.

The final research question is *what factors influence the adoption of big data by Dutch consultancy SMEs?* This empirical question will be answered with the use of interviews. This enables a rich understanding of the relationships between the challenges that SMEs face and the adoption of big data.

1.5 Outline

This thesis is structured as follows. The next chapter discusses the concept of big data and provides a definition, thereby answering the first research question.

This is followed by chapter 3, which discusses the resource-based view. Together with dynamic capabilities and the preceding chapter, it conceptualizes big data as a unique resource. This answers the theoretical part of the second research question.

Chapter 4 contains the conceptual section of this thesis. A conceptual framework is constructed, based on the literature review of Sun, Cegielski, Jia, & Hall (2018). This presents an overview of the most studied factors of influence in the use of big data. These are considered the challenges that SMEs face and therefore provides the answer to research question three.

Chapter 5 discusses the methods used for the empirical part of this thesis. The constructed conceptual model is used as a basis to guide the interview questions and scope the research. The chapter begins by discussing the research strategy and design, followed by the methods of data collection and analysis. The final section of the chapter discusses how the quality of the research was ensured.

Chapter 6 presents the first results from the interviews. It discusses whether the interviewees use big data according to the definition of this thesis. Then it briefly explains what use cases were found. By doing so, this chapter provides the answer to the empirical part of research question two.

Chapter 7 will present the majority of the interview results, as it answers the focus of this research with research question four. It will discuss which factors were found to be the most challenging for the adoption of big data and how SMEs have leveraged their resources to overcome them.

The final chapter consists of the conclusion and discussion of this thesis. It summarizes the findings, provides conclusions and formulates recommendations. The main research question is answered, which is followed by a discussion of the limitations of the research and directions for future research.

Chapter 2 – Big data

This chapter will provide context to the research by discussing big data. Big data is often proclaimed to be a new source of competitive advantage, but is poorly defined (McAfee & Brynjolfsson, 2012; Porter & Heppelmann, 2015; Raguseo, 2018; Wamba et al., 2017). Multiple definitions and interpretations exist and discussing these is essential for understanding what big data is. This chapter provides an answer to the first research question: *what is big data?*

2.1 Defining big data

In the last decade, both the capacity for storing and processing data have increased (Russom, 2011). At the same time, the price to store and process data fell. This resulted in a situation in which it was no longer costly to gather, store and analyze data. The threshold to store data was lowered, which opened the door to storing different formats of data (Cavanillas et al., 2016). Prior to these developments, most of the data of firms was stored in relational databases or spreadsheets (Hammer et al., 2017). This data is consistent, orderly and easy to represent in tables (Simon, 2013). Tables could be linked together, creating a relational database. However, information can be found in other formats as well. Now it is viable to store unstructured data like text, audio or video (Cavanillas et al., 2016). This enabled the exploitation of data about things that were previously much more difficult to store and analyze.

A process of datafication was set in motion, in which everything can be converted to data (Cukier & Mayer-Schoenberger, 2013). At the same time, the perceived value of data has increased. The philosophy that you can improve something if you can measure it is widespread and data has often been called the new oil (Hartmann et al., 2016). Data might contain hidden information and it is only a matter of processing it to get to new insights. The thinking goes that the more data you can analyze, the more insights you can extract.

The result is an increase in the volume and variety of data, but also a higher velocity in which data is gathered. While large volumes of data are now cheap and easy to store, there is a bottleneck. Extracting information from these large – and fast growing – volumes of data requires a lot of processing power. Certainly, processing power has increased tremendously, but not at the same rate as storage capacity (Russom, 2011). Data is gathered faster than it can be analyzed. All these developments have led to a new concept: *big data*.

The first mention of the term big data in a scientific article was in 1997 (Press, 2013). The article in question discussed the problem that data created when it became too big to handle by the RAM on a computer or when it did not fit entirely on a storage disk. This was the time in which a computer with a storage capacity of over 2GB was considered overkill (Miastkowski, 1998). Nowadays, this storage capacity is insufficient even for smartphones. This indicates that the problem of big data could be dependent on time and context.

Herein lies the problem of defining big data. There is no consensus on how large the volume of data should be, how fast data should be generated or how many types of data formats should be collected before data is 'big' (Cavanillas et al., 2016). Nevertheless, these characteristics are useful in discerning the differences between regular and big data. The volume, velocity and variety are the three characteristics introduced in a 2001 paper of the Meta Group that together form the 3V model (Laney, 2001).

Figure 1 shows the 3V model and increasing complexity along each of the axes. Note that it does not mean that big data has to be a certain amount of terabytes in size with real time data collection of unstructured data like photos. Instead, the 3V model is used as an overview of the most important characteristics of big data. The definition this research will use is based on Stonebraker (2012), which incorporates only these three characteristics: *big data is either high volume, high velocity or high variety information assets*. This definition stays true to the core of what big data is and enables a combination of the core characteristics.

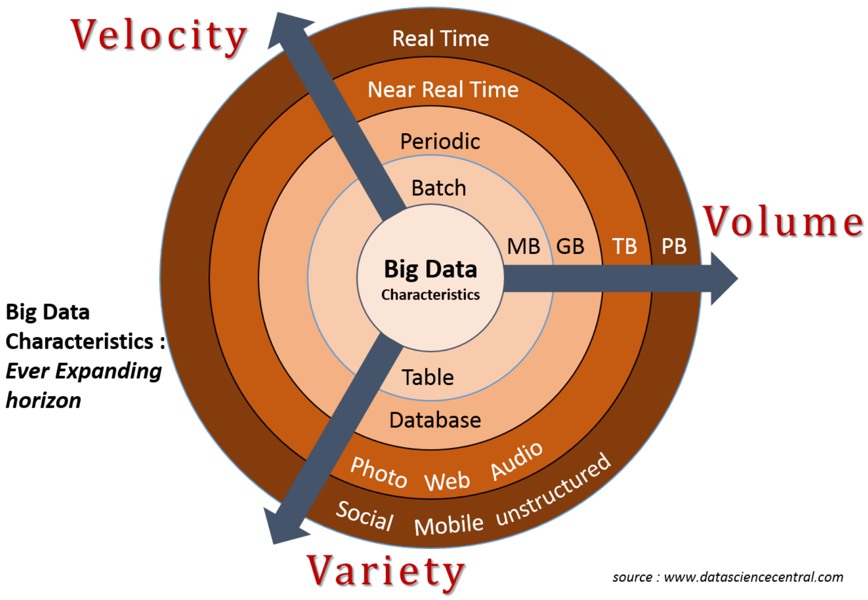

Figure 1 - 3V model of Big Data (Anwar & Khaki, 2017)

It is important to acknowledge that other characteristics have been proposed to define big data, as this frustrates the process of settling on a definition. Characteristics like veracity, value, volatility and many more have been included in subsequent iterations of the 3V model (Cavanillas et al., 2016; Hammer et al., 2017). These characteristics describe what big data can do, but not what it is (Gandomi & Haider, 2015).

Another, frequently mentioned characteristic of big data is the need for new database management tools (Cavanillas et al., 2016). These new tools are needed due to the problems that arise in the processing of big data. However, new soft- and hardware has been developed to cope with these issues. Technologies like cloud computing and techniques like big data analytics (BDA) could tackle the issues posed by big data (Manyika et al., 2011).

The problem in incorporating the need for specific technologies into the definition of big data is twofold. First, as previously mentioned, 'big' is a fluid concept that depends on time and context. What was considered big ten years ago, might not be anymore today. In addition, the development of soft- and hardware has resulted in new technologies specifically designed for big data, but older technologies have been adapted as well (Manyika et al., 2011). A definition like "*big data is a term encompassing the use of techniques to capture, process, analyze and visualize potentially large datasets in a reasonable timeframe not accessible to standard IT technologies (...)* " (NESSI, 2012, p. 6) opens itself for interpretation and discussion about the used technologies, rather than the data itself. Specialized technologies could be used as an indicator for big data processing, but are not the same as big data.

This leads to the second reason to exclude technologies from the definition of big data. The technologies used to gather, store and analyze big data are simply tools. They are necessary to extract value, but are not part of big data. For example, you would not define the data in a spreadsheet by a tool like Microsoft Excel to visualize it.

2.2 Summary

This chapter briefly discussed the concept of big data and provided a definition to guide the thesis. The chapter started with explaining the developments that gave rise to the term big data. It is a popular term, but lacks consensus on an exact definition. The characteristics of big data should inform the definition, but following the original three characteristics – volume, velocity and variety – many more have been proposed. The fuzziness of the definition is not helped by the fact that some have included aspects related to its use, like the value, or the method by which it is utilized, like BDA. In this thesis, the definition of big data will stick to the core of what it is: information assets that are either high in volume, velocity or variety. In doing so, the first research question is answered: *what is big data?* The next chapter discusses why big data is a unique resource.

Chapter 3 – Big data as a source of competitive advantage

The previous chapter provided a definition of big data. This chapter will discuss the resource based view (RBV) and dynamic capabilities (DC) to explain why big data is a unique resource. This will frame big data in the overall discussion of the link between IT and firm performance, and provides a foundation for the thesis in scientific literature. It also highlights the contribution of this research. As a result, this chapter provides an answer to the theoretical part of the second research question: *why is big data a unique resource and how does it differ from traditional IT in Dutch consultancy SMEs?*

3.1 Unique resources in the RBV

A theoretical lens is needed to understand the importance of big data for SMEs. In the literature it is claimed that it could lead to a performance boost or competitive advantage (Hammer et al., 2017; McAfee & Brynjolfsson, 2012), but the theoretical relationship has to be explained. This section will discuss the resource-based view (RBV) and how resources are conceptualized. The next section discusses dynamic capabilities (DC) to exploit these resources. Together, these perspectives offer an explanation for the performance of firms and what role big data might play in creating a competitive advantage.

First, it is important to understand what organizational resources are. A 1991 publication on the RBV by Jay Barney laid the groundwork for what later would become the resource-based theory (RBT) (Barney, Ketchen, & Wright, 2011). This theory focused on the resources of a specific firm to explain the creation and sustaining of a competitive advantage (Priem & Butler, 2001). This thinking was new because strategic management had mostly been focused on external factors like the industry structure of Porter (1985).

In the RBV, firms and their (access to) resources take a much more prominent role in explaining competitive advantages (Barney et al., 2011). Not only do the resource of a firm matter, but also its capabilities to exploit them (Teece et al., 1997). In other words, a competitive advantage can be created by a combination of resources and capabilities. While some critics questioned the usability of the RBV in a management setting (Priem & Butler, 2001), it has developed to one of the most prominent theories on understanding organizations (Barney et al., 2011).

In the RBV, firms are conceptualized as bundles of resources, which can be heterogeneously distributed across firms (Eisenhardt & Martin, 2000). Some firms might have a larger bundle of resources than others in the same industry. These differences between firms can persist over time, creating certain path dependencies (Teece et al., 1997). In addition, the transfer of resources comes at a cost, making resources ‘sticky’ (Priem & Butler, 2001). Resources are therefore key in formulating strategies that create value (Eisenhardt & Martin, 2000).

Resources can be tangible assets like capital, equipment and geographical location or intangible assets like technological or managerial know-how, experience and learning (Newbert, 2008). There are four characteristics that determine, together with the capabilities of a firm, the likelihood of creating and sustaining a competitive advantage. These four characteristics are value, rareness, imitability and non-substitutability (VRIN) (Kraaijenbrink, Spender, & Groen, 2011).

A resource has value if it contributes to the efficiency or effectiveness of a firm (Priem & Butler, 2001). Think of manufacturing equipment that increases overall production of a firm. The second characteristic is the rareness of a resource. Scarce resources have a higher rareness, which could increase their worth (Priem & Butler, 2001). An example would be the mismatch of the few available skilled data scientists and current high demand. These two characteristics form the basis of creating a competitive advantage, but the latter two are instrumental in sustaining one.

If a resource is easy to imitate, replication by competitors is easy and the firm might find it difficult to sustain a competitive advantage based on that specific resource (Priem & Butler, 2001). Tangible resources are easier targets for imitation than intangible assets like tacit knowledge (Teece et al., 1997). The final characteristic is about the substitutability of the resource. If other resources can be used to fulfill the same function, sustaining a competitive advantage is more difficult as well (Priem & Butler, 2001).

Thus, the VRIN characteristics of organizational resources determine how unique it is and influence the competitive strategy and subsequent competitive advantage of firms. Firms with the same resource endowment might still perform differently due to their different routines and dynamic capabilities (Eisenhardt & Martin, 2000). I will briefly discuss dynamic capabilities in the next section.

It is important to be aware of the critiques the RBV has to apply it correctly. It is accused of lacking any managerial implications, as it does not indicate *how* a firm should acquire resources or develop capabilities (Kraaijenbrink et al., 2011). However, the RBV was never meant to help managers guide their companies, but rather as a way of explaining differences in performance between firms.

Another important critique is the tautological nature of the RBV. Both the definition of value – "*resources are valuable when they enable a firm to conceive of or implement strategies that improve its efficiency or effectiveness"* – and resources – "*resources will be defined as stocks of available factors that are owned or controlled by the firm*" – are tautological (Kraaijenbrink et al., 2011). The problem that arises as a result of these tautologies is that the RBV will always be correct. A solution could be to be more specific about the types of value and resources. For example, there is no distinction between resources: tangible-intangible, human-technologic, static or dynamic. By making clear distinctions, it is easier to understand how each resource plays a role in contributing to firm performance (Kraaijenbrink et al., 2011).

Furthermore, in itself the RBV is mostly focused on the internal workings of the firm. It theorizes that firms pick and enter markets where they will get the largest rents on their unique resources (Teece et al., 1997). However, not all firms might be able to identify their own unique resources and act accordingly. This also neglects external factors that could influence the ability of the firm to create a competitive advantage. Combining the RBV with dynamic capabilities solves this problem, as the latter is not solely concerned with internal processes and considers external factors as well.

In this thesis, the RBV is used as a theoretical lens to understand the relationship between organizational resources and firm performance. By complementing it with dynamic capabilities While it might not have direct managerial implications, using the RBV in a specific context could still yield indicative explanations (Kraaijenbrink et al., 2011). Furthermore, by studying specific resources like big data, the tautological nature of the RBV is avoided.

3.2 Dynamic capabilities

The RBV is useful in conceptualizing firms, but cannot fully explain the differences in performance between firms with the same resource endowment (Shan, Luo, Zhou, & Wei, 2018). Having organizational resources is only useful if a firm has the appropriate routines to exploit them. When a firm does not have the appropriate routines, it could try to change them (Easterby-Smith & Prieto, 2008). The changing of routines and organizational resources is referred to as dynamic capabilities.

In a broader sense, the dynamic capabilities of a firm refer to the processes or routines of a firm to exploit resources (Teece et al., 1997). Dynamic capabilities are defined as the "*firm's ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments*" (Teece et al., 1997, p.516). Thus, the better a firm is able to acquire, integrate, exploit, shed and recombine resources, the more likely it is to create a competitive advantage (Eisenhardt & Martin, 2000).

Dynamic capabilities are essential for adapting in changing environments. A firm has routines in place to create value from its resources (Teece et al., 1997). When the environment changes, a particular routine or resource could become suboptimal (Shan et al., 2018). The ability to change these routines is what makes DC powerful. However, path dependencies and historical organizational assets do impact the choices and options available to a firm. If a firm chose to build competences in a certain area in the past, then future options are dictated by this decision (Teece et al., 1997).

The process to change routines is dependent on the ability of the firm, and its people, to sense opportunities (Teece et al., 1997). Without this ability to identify possibilities, not much will change. The second step is to seize the opportunity and commit time and resources to the change. This is dependent on the strategic insight and willingness of managers. However, path dependencies might offer fewer possibilities than managers realize (Teece et al., 1997). The change has to be internalized in the firm for it to take effect. A routine can only change when it is the accepted cultural norm of the firm for example.

A similar collective effort is the learning process within organizations (Teece et al., 1997). Learning enables processes to be more effective or efficient and could increase the chances of recognizing opportunities. Learning can happen at the individual level, but when knowledge is shared between members of the firm, it can be captured as organizational knowledge in the form of routines (Teece et al., 1997).

Combining the RBV and DC enables the application of an internally focused theory of the firm, while leaving room for external factors (Shan et al., 2018). The dynamic capabilities of a firm are especially useful in explaining the acquisition and integration of external resources into the firm. Within the DC theory, there is more room for strategic management and decision-making (Teece et al., 1997). The 'dynamic' part of DC refers to the capacity of a firm to change in order to align with changing business environments (Teece et al., 1997).

As such, the DC is seen as an extension of the RBV (Easterby-Smith & Prieto, 2008). It has similar critiques, for example in its tautological nature of defining 'capability' as 'ability' (Easterby-Smith & Prieto, 2008). It is also claimed to be tautological due to the fact that studies focus on sustained competitive advantages, which are then attributed to dynamic capabilities (Easterby-Smith & Prieto, 2008; Eisenhardt & Martin, 2000). Clarifying that capabilities refer to the potential to do something has solved the first critique (Easterby-Smith & Prieto, 2008). The second critique is based mostly on the lack of empirical evidence in dynamic markets (Eisenhardt & Martin, 2000), an area in which this research can contribute.

3.3 Big data as a source of competitive advantage

In the previous sections the concepts of big data, organizational resources and dynamic capabilities were discussed. This section will bring these concepts together to answer the theoretical part of the second research question: *why is big data a unique resource and how does it differ from traditional IT in Dutch consultancy SMEs?* It is necessary to answer this question, to fully understand how big data could play a role in increasing productivity and fostering a competitive advantage.

The characteristics of organizational resources that play a role in determining the potential for a competitive advantage are the resource's rareness, imitability, non-substitutability and value (Kraaijenbrink et al., 2011). If big data possesses these VRIN characteristics, it has a high potential for being a distinct resource that firms can exploit. However, a firm needs to have appropriate DC in order to adapt and adequately exploit these resources. DC are the link between the organizational resources of a firm and the ability to adapt and survive in a changing environment (Teece et al., 1997).

To assess whether big data is a unique resource, its characteristics are important. The characteristics of big data state that it consists of information assets high in volume, velocity or variety (Stonebraker, 2012). In other words, these are specific information assets that are part of the organizational resources of a firm. A firm in the possession of big data can be seen as rare, as the use among SMEs is lagging behind (Coleman et al., 2016). Another aspect of the rareness of big data refers to what information the data captures.

To illustrate, firms like Facebook are currently leading in the capturing of big data (McAfee & Brynjolfsson, 2012). Due to its millions of users, Facebook possesses a uniquely detailed database with information on visitors of the platform. It uses its data to tailor advertisements to specific target groups, thereby creating a value proposition for advertisers. While data on people might not be rare in itself, the fact that a firm possesses big data can lead to new insights that were previously hidden (NESSI, 2012). The volume of the data that Facebook possesses makes it a rare resource, as very few other companies possess such detailed information advertisers can use. In the case of consultancy firms, the data available to them might enable them to solve unique problems posed by their clients. This will increase the worth and reliability of their services (Wamba, Akter, Edwards, Chopin, & Gnanzou, 2015).

The second characteristic refers to the ability to imitate organizational resources (Kraaijenbrink et al., 2011). Imitating the organizational resources of competitors is already difficult, due to the coherence between resources and competences (Teece et al., 1997). Big data is more complex due to its high volume, velocity or variety, which increases the cost of imitation. Furthermore, data could be proprietary, meaning that the owner of the data is the sole entity with access to it. In the consultancy industry, firm could either have big data themselves or utilize data of their clients. In the first case, they could sell BDA and advice based on their own data (Ardagna, Ceravolo, & Damiani, 2016). This results in big data being difficult to imitate. In the second case, it is dependent on how many other parties have access to the data.

The third characteristic is non-substitutability (Kraaijenbrink et al., 2011). Due to the characteristics of big data, it is also hard for others to substitute big data with big data of their own. This refers back to what is captured within the data, but also to the fact that larger volumes of data could lead to higher value (Cavanillas et al., 2016). For example, firms competing with each other in the industry of targeted advertising could distinguish themselves with big data. If one firm has simply more data available about a certain market segment, it is better suited to sell advertisements as it can make more accurate predictions about the needs and characteristics of that segment (Manyika et al., 2011; NESSI, 2012). In this sense, it is very hard to substitute big data for competitors with something that might give the same level of insight and predictions.

The final characteristic is about the value of the organizational resource (Kraaijenbrink et al., 2011). This is in large part determined what big data is used for. It can be used for decision-making, increasing efficiency in manufacturing processes or distinguishing yourself from the competition by offering more reliable, data based services (Manyika et al., 2011). However, it has to represent value for the firm (Wamba et al., 2015). This is where the fuzziness of big data surfaces. It is very broad and has many different applications (Manyika et al., 2011). Depending on the situation, it might be uniquely suited to represent value, or the same value might be easily achieved with a different resource. Nevertheless, this research does assume that big data represents value.

For example, analyzing raw data with BDA has a high potential for optimizing business efficiency and effectiveness, both operationally and strategically (Ardagna et al., 2016; Wamba et al., 2017). Extracting hidden insights from data to inform decision-making and better understand processes is potentially valuable to firms. Data itself could even be considered as a product, which has given rise to Big Data Analytics as a service (BDAaaS) (Ardagna et al., 2016). One could argue that the strength of big data is in the extraction of hidden correlations and insights from data without a predefined goal (Cukier & Mayer-Schoenberger, 2013). In this sense, the use cases of big data are only limited by the amount of data itself and the imagination for developing new value propositions. Therefore, big data can be considered as a unique resource in the RBV.

However, the value of a resource is also determined by the capability of a firm to exploit it (Teece et al., 1997). Big data could play a significant role here, not only as a resource but as a capability-builder (Wamba et al., 2017). For example, big data could foster learning, thereby increasing the development of organizational knowledge and dynamic capabilities. This is illustrated in Figure 2, which shows the transformation from raw data as a commodity to a distinct resource (Roden, Nucciarelli, Li, & Graham, 2017).

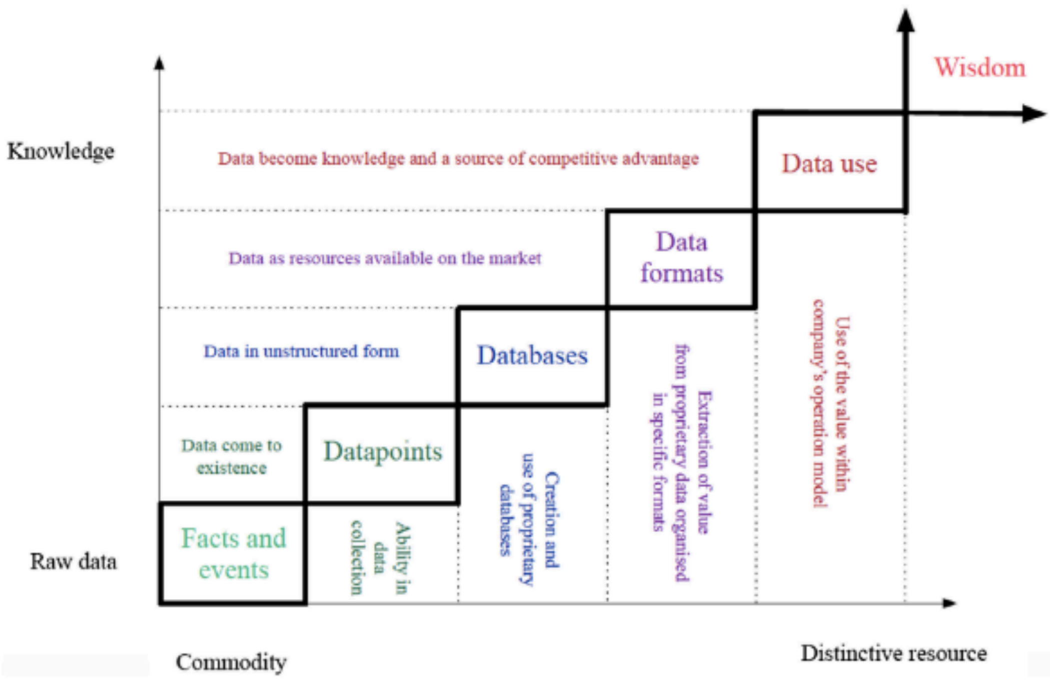

Figure 2 - The data ladder (Roden et al., 2017)

The steps along the way indicate the impact on and integration of data within the firm. Eventually, data leads to the development of higher-order capabilities and knowledge that increase the ability of the firm to create a competitive advantage (Kraaijenbrink et al., 2011). This represents both learning from data and the value of applying the data. This could create a positive feedback loop, in which learning from data increases the IT capabilities of a firm, while enhancing the potential for a competitive advantage.

To better illustrate how the previous sections are related, the proposed research model is shown in Figure 3. First, the characteristics of big data make it a VRIN organizational resource, which could influence the competitive advantage of a firm. However, big data can also influence the building of dynamic capabilities through learning (Roden et al., 2017). In turn, dynamic capabilities enable a firm to change existing routines and better exploit organizational resources (Eisenhardt & Martin, 2000). In this view, both the RBV and DC are necessary for a firm to create a competitive advantage. However, the RBV is mostly internally focused and does a poor job of considering external influences (Priem & Butler, 2001). In the proposed model, these external factors influence the DC of a firm. Therefore, a firm is confronted with decisions and trade-offs resulting from internal and external pressure, to better represent reality (Easterby-Smith & Prieto, 2008).

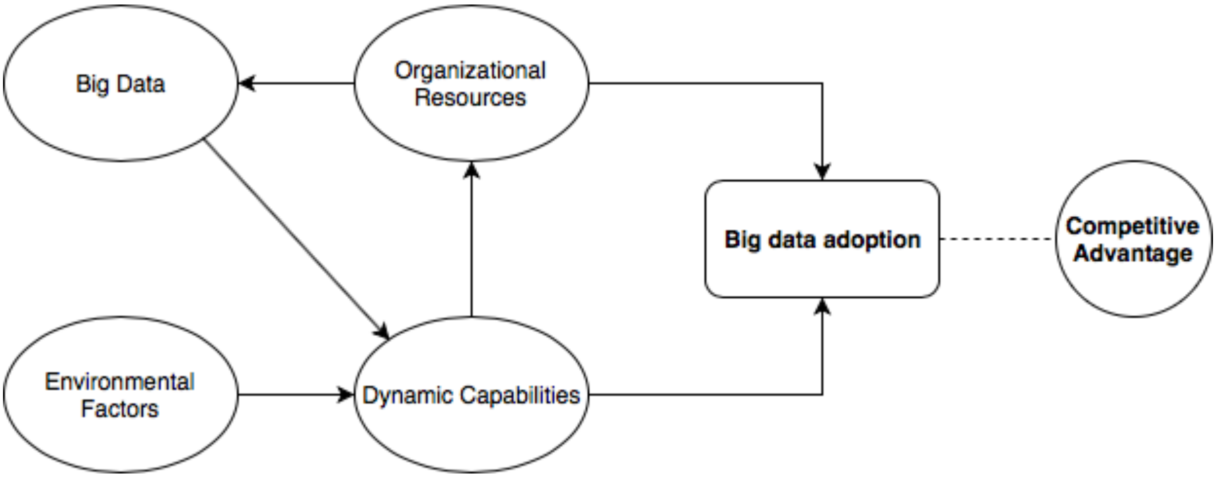

Figure 3 - Proposed research model

The proposed model enables this research to focus on the internal workings of a firm, but also incorporate external pressures that might be encountered in the adoption of big data. For example, new legislation could force firms to adapt and reconfigure existing routines, resources or competences to cope with the new reality (Priem & Butler, 2001). This is not only essential to avoid outdated or suboptimal routines in the creation of a competitive advantage, but for firm survival in general (Lefebvre, Lefebvre, & Harvey, 1996; Zahra, Sapienza, & Davidsson, 2006). The focus of this research will be on the relationships between big data adoption, organizational resources, dynamic capabilities and environmental factors. It is assumed that these contribute to a competitive advantage, but as shown in Figure 3, it is not part of the relationships investigated here.

This is due to the fact there is some discussion on the role of big data. While some academics and practitioners are hailing big data as a new source of competitive advantage, innovation and decision making (Coleman et al., 2016; Manyika et al., 2011; H. G. Miller & Mork, 2013), others are more skeptical (Wamba et al., 2017). If big data is truly transformational, it will pose challenges for firms to adopt it. Dynamic capabilities are then necessary to explain how these firms are able to do so.

However, if big data is framed in the broader context of investment in information systems, results on the increase of firm performance are mixed (Wamba et al., 2017). Some have found positive influence (Gunasekaran et al., 2017), but others did not see an increase in the performance of firms after the investment in information systems that increase efficiency (Chae, Koh, & Prybutok, 2014).

A possible explanation could be the time lag between investment and performance increase, making it more difficult to pinpoint what exactly was responsible for the increased performance of the firm in the first place (Wamba et al., 2017). By investigating the role of big data in relation to organizational resources and dynamic capabilities, a better understanding can be created of how it might indirectly influence firm performance or create competitive advantages.

3.4 Summary

This chapter has answered research question two: *why is big data a unique resource and how does it differ from traditional IT in Dutch consultancy SMEs?* The chapter started with a discussion on the theoretical lens used to understand how big data might be involved in creating a competitive advantage. The RBV conceptualizes firms as bundles of resources and the capability to exploit them. A competitive advantage might be created and sustained by the resource endowment of a firm and how well it is able to exploit its bundles of resources. Resources come in many different forms and can be tangible or intangible. The more difficult it is to imitate or substitute a resource, the better it is suited for the creation of a competitive advantage. The value and rareness of a resource also play a role in its distinctiveness. Thus, truly unique resources are valuable, rare and difficult to imitate or substitute.

In addition, the dynamic capabilities of a firm enable it to integrate, build and reconfigure internal and external competences in response to changing environments. The better the DC of a firm, the likelier it is to successfully adapt and exploit organizational resources. The ability to change existing routines and internalize new knowledge or competences is essential in changing environments. Firms do not operate in a vacuum and can receive external incentives. Together with the RBV, DC explain how firms cope with internal and external influences based on their history, competences and assets.

The final section of this chapter discussed how big data could be considered a unique resource based on the VRIN characteristics. Big data is difficult to imitate and substitute and might be considered rare and valuable. It could very well lead to competitive advantage on its own, but in this research it is assumed that these characteristics alone do not make it sufficiently different from traditional IT investments. Rather, the characteristics of big data make it more suitable for the development of dynamic capabilities, enabling firms to better leverage existing resources and big data to create a competitive advantage. The real uniqueness might not be in the characteristics of big data, but in the creation of knowledge and capabilities that mediate between big data and firm performance.

The empirical part of the second research question will be answered in chapter six. The next chapter will focus on the challenges of big data adoption in order to answer the third research question.

Chapter 4 – Challenges of big data adoption

This chapter will construct and discuss a conceptual framework based on a literature review and relevant concepts of the RBV. By looking at the factors influencing the adoption of big data, challenges can be distilled. These can be related to the organizational resources, big data itself or the environment of the firm. This answers the third research: *what are the potential challenges of big data adoption?* The conceptual framework will be applied for the empirical part of this research in chapter 7.

4.1 Factors of influence

The starting point for the conceptual framework is to look at the established body of knowledge on the adoption and use of big data. Systematic literature reviews provide an overview of the current knowledge by summarizing findings of other studies (Creswell, 2014). To find literature reviews on the topic of big data use, the search started with the terms *big data use* and *big data adoption* in the Scopus and Google Scholar databases. One of the results was an article by Sun, Cegielski, Jia, & Hall (2018), who performed a literature review to index the most often studied factors influencing the adoption of big data by organizations. This resulted in a framework consisting of 26 factors based on 62 articles (see Figure 11 in Appendix A).

The framework does not indicate whether these factors have a positive or negative influence, nor does it mean that factors with a higher frequency are more important. It does showcase the current focus of the academic literature. This literature review will be used as the conceptual framework, as the mentioned factors are considered to be of influence on the adoption of big data and indicate possible challenges.

Sun et al. (2018) grounded their literature review in the Diffusion of Innovation (DOI) theory of Rogers (2003) and the Technology-Organization-Environment (TOE) framework. Together these concepts explain how innovations spread through populations and what factors influence the decision to adopt and use a specific technology. These factors can be categorized in one of the three contexts of the TOE: Innovation Characteristics (IC), Organizational Characteristics (OC) and Environmental Characteristics (EC).

These characteristics fall under certain concepts of the proposed research model. For example, IC falls under big data, EC under environmental factors and OC under both organizational resources and dynamic capabilities. The updated research model is shown in Figure 4 and provides an overview of the important concepts and their relations studied in this research. The next section will briefly discuss the factors per category from the literature review of Sun et al. (2018).

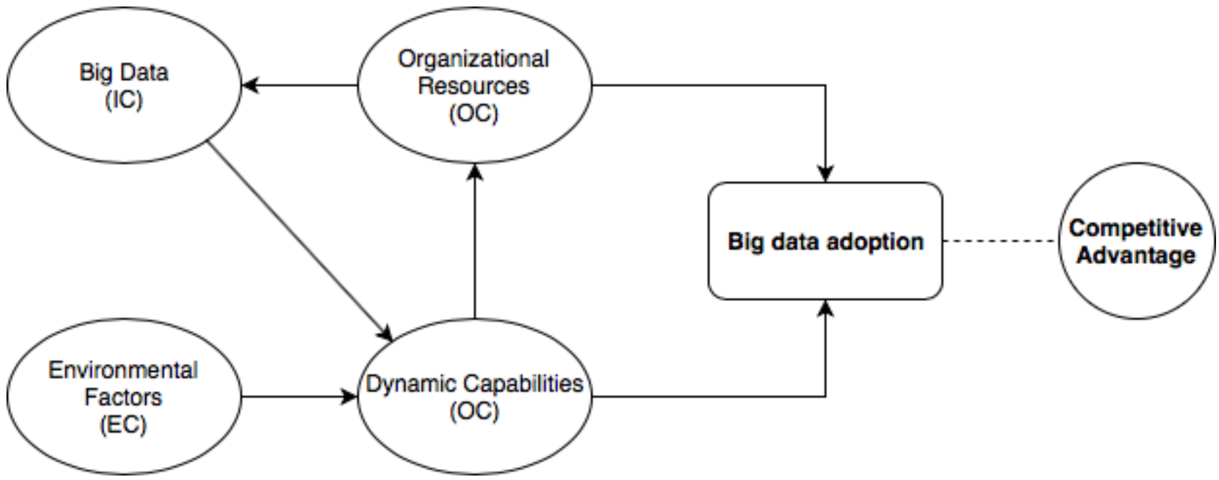

Figure 4 - Updated research model

4.1.1 Innovation Characteristics

The focus of this thesis will be on the organizational and environmental factors, due to the use of the RBV and DC as theoretical lens. Furthermore, big data might differ wildly per firm in what it is and how it is used. As a result, the IC factors are considered to be part of the technological context rather than the actual object of study. Big data itself is not what is being investigated, but how firms cope with their resources and capabilities in order to adopt big data. The IC could very well play an important part in influencing the adoption or use of big data by SMEs, but they are not part of the main relationships investigated here.

To illustrate, the IC category consists of six factors in the Sun et al. (2018) literature review. Five factors are from the DOI of theory of Rogers (2003) These determinants affect the rate of adoption throughout a population. The first is relative advantage, which refers to the perception of how much better an innovation is compared to current tools or techniques (Rogers, 2003). The second is compatibility of the innovation with the existing systems that are in place. The third characteristic is the complexity or difficulty to learn or master an innovation. Trialability is the fourth and refers to the opportunity and ease to test out the innovation before committing. The final characteristic observability is about others might perceive the results of the innovation. These characteristics are interdependent and therefore often considered as a whole for a particular type of innovation (Rogers, 2003). Sun et al. (2018) added the cost of adoption as a sixth factor. The price for acquiring an innovation could be considered part of its characteristics as well.

4.1.2 Organizational Characteristics

The category Organizational Characteristics notes all the factors that relate to the user or adopter of an innovation. In this research the unit of analysis is the SME. More specifically, these factors are the main focus of this thesis as they contain the organizational resources and dynamic capabilities. Sun et al. (2018) categorized 12 factors under the OC category. These and their suspected roles will be briefly explained below.

The role of **human resources** has been studied extensively in the adoption and consequences of new technologies (Antonioli & Della Torre, 2016; Benhabib & Spiegel, 2005; Freel, 2005; Ramanathan, Philpott, Duan, & Cao, 2017; Wozniak, 1987). Human resources consist of the people and their knowledge and skills that are relevant to the organization. If the uncertainty of new technologies is high, the adopter has to rely on the skill of its human resources (Wozniak, 1987). Furthermore, the training of human resources have been found to positively affect the performance of firms after technology adoption (Boothby, Dufour, & Tang, 2010).

In this sense, a human resource as a factor is both part of the assets of the firm – its people, skill and knowledge – and its ability to change, thus a dynamic capability. Specific to big data, new skills might be needed to interpret and analyze data and to recognize opportunities. This could manifest itself in the need for data scientists and specific IT human assets (Al-Qirim et al., 2017; Ramanathan et al., 2017). Currently, a skills gap of specialized personnel in Europe and The Netherlands means that data scientists are more expensive to hire (Cavanillas et al., 2016). The expected challenge is therefore that a lack of appropriate human resources hinders the use of big data by SMEs.

The **technology resources** of the firm refer to the quality and infrastructure of the IT systems in place (Sun et al., 2018). A study found that the quality of the existing information systems influences the decision to adopt newer IT systems in SMEs (Rashid & Al-Qirim, 2001). The acquisition of new IT is therefore dependent on the appropriate in-house resources and the IT already in use (Fink, 1998). This is related to the compatibility of the current systems, as this has impact on the ease of integration of newer technologies (Ramanathan et al., 2017). Therefore, without the proper technology resources, the use of big data by SMEs is challenged.

If resources are available, it still depends on the management to allocate sufficient resources for the use of big data. Without the **management support**, new and often risky or uncertain projects do not receive the needed resources for successful acquisition (Alshamaila, Papagiannidis, & Li, 2013). Prior studies found that resistance (Rashid & Al-Qirim, 2001) and IT experience (Fink, 1998) of management influence the adoption of IT systems. One study found that if top management did not realize the value of big data analytics, it did not adopt it (Verma & Sekhar, 2017). Furthermore, the management plays an important role in strategic management of the firm and facilitating the dynamic capabilities (Teece et al., 1997).

Technology readiness refers to the internal expertise and infrastructure of the firm to adequately use big data (Sun et al., 2018). In related literature it refers to the alignment of a technological solution with the firm's strategy (Rashid & Al-Qirim, 2001), but it is mostly about the level of expertise of the personnel (Verma & Sekhar, 2017) and the flexibility of the

IT infrastructure (Wamba et al., 2017). In this research it is considered to incorporate the factors **human resources** and **technology resources**.

The factor **organization / IT structure** is about how well the organization and IT structures are developed (Sun et al., 2018). It also incorporates how easy it is to collaborate within the organization and the configuration of the IT department and personnel. This touches upon aspects of the **technology readiness**. Specifically to SMEs, it is unclear how important the organizational structure is. With particularly small firms, there might not even be a distinction between centralized or distributed functions and relationships, thereby rendering them irrelevant (Coleman et al., 2016). Therefore, this factor is excluded from the framework.

The **decision-making culture** of a firm is about how decisions are made. Whether there are certain norms in place, like basing decisions on evidence, influence the culture of decision-making at the firm level (Sun et al., 2018). Within SMEs, the power for making decisions is often concentrated in the owner or founder of the firm (Hausman, 2005). As a result, the innovativeness of the founder is largely responsible for the overall innovativeness of the firm (Hausman, 2005). Big data can play a role in transiting to data-based decision-making, but the management or owner of a firm have to accept this (Al-Qirim et al., 2017). If the founder is risk-averse, the likelihood of using big data by SMEs is therefore lower.

Similarly, the **business strategy orientation** is about incorporating big data within the strategy of the firm (Sun et al., 2018). It is about how IT plays a role in the market selection, development and diversification (Levy, Powell, & Worrall, 2005). These are choices that the management has to make, which can be based on the integration of IT and data within the organizational strategy (Ramanathan et al., 2017). If the business strategy is archaic, then it is likely to be centralized and not data-driven (Caesarius & Hohenthal, 2018). The business strategy is dependent on the business understanding of the management (Verma & Sekhar, 2017). As a result, it is related to **management support** and **decision-making** culture.

The factor **business resources** is defined as the availability of the resources that are adequate for adopting big data (Sun et al., 2018). It is about the sharing of information between departments and the policy that is set towards adopting big data. According to Wamba et al. (2017), business resources can be considered to overarch big data infrastructure, management capabilities and big data expertise of employees. As a result, the factor incorporates aspects of **human resources**, **technology resources** and **management support.** Due to this overlap this specific factor is excluded from the conceptual framework.

The factor **change efficacy** is about the ease with which employees embrace new technologies within the organization (Sun et al., 2018). The more efficient an employee is in coping with changes, the likelier the use of big data. Positive attitudes of employees are also found to influence the successful integration of new technologies (Fink, 1998). Within the literature, the flexibility of an organization is found to play a significant role in the decision to adopt new technologies like big data (Raymond & Uwizeyemungu, 2007). Furthermore, SMEs are thought to be more flexible due to their size. This factor is essentially a dynamic capability, as it is about the ability to integrate external competences into the firm. Therefore, the expected effect is that the more flexible the SME is, the likelier it will be to use big data.

The factor **Information Systems (IS) strategy orientation** is related to that of **business strategy orientation**. If a firm prioritizes the use of information systems, it will be more open towards using big data (Sun et al., 2018). However, without this priority setting, it will probably wait until market signals force firms to adopt (Levy et al., 2005). Therefore, firms that already prioritize data use, for example from external sources, are more likely to adopt big data (Kwon, Lee, & Shin, 2014). The willingness to explore options related to big data and incorporate these into the strategy of the firm will positively influence the use of big data (Caesarius & Hohenthal, 2018).

In most statistics, the **firm size** is determined by the number of employees (DG GROW, 2018). While it is somewhat arbitrary, a firm with 250 employees or fewer is considered to be an SME. Table 1 shows the different categories of SMEs per headcount, or less frequently used, the turnover of the firm.

Table 1 - Definitions of different SME categories (DG GROW, 2018)

Company category	Staff Headcount	Turnover or balance sheet total
Medium-sized	<250	≤ €50m or ≤ €43m
Small	<50	≤ €10m or ≤ €10m
Micro	<10	≤ €2m or ≤ €2m

The larger a firm is, the more resources it has at its disposal to direct towards the acquisition and use of big data (Sun et al., 2018). While an organization might become less flexible the larger it gets, one study found that SME size did not play a role in the capacity to adopt new IT innovations (Gray, 2006). Indeed, IT has often been characterized as leveling the playing field between large and small firms (Alshamaila et al., 2013). However, due to more resources being available, larger firms are better able to absorb risk and uncertainty related to big data (Agrawal, 2015). This research excludes this factor as it is part of the context within with the study is carried out.

The final organizational characteristic is **appropriateness**, or the timing of the use of big data (Sun et al., 2018). When big data is utilized during an appropriate time, it could increase the performance of the firm. However, before a technology is used, a whole process has preceded it. Using a technology is thus not instantaneous, but rather a process (Rogers, 2003). A study found that the performance of firms was negatively impacted due to the time lag between the availability of new technologies and its adoption by firms (Hoppe, 2002). However, this thesis is studying the process before and during the use of the technology, not the impact of the technology on the performance of firms. Therefore, this factor is excluded from the framework.

4.1.3 Environmental Characteristics

The remaining 8 factors fall under the environmental characteristics category. These factors are external to the firm and are theorized to influence its dynamic capabilities. This section will briefly discuss all EC factors, starting with the factor **security, privacy and ethics concerns regarding collecting data**. It is about the legal and ethical concerns that might arise in collecting data, for example on customers (Sun et al., 2018). So far, the empirical literature has focused on security issues about storing data (Coleman et al., 2016). Regarding the security of data, one study found that SMEs are driven to adopt cloud computing for the added security (Gupta, Seetharaman, & Raj, 2013). While ethics have often been discussed (Michael & Miller, 2013; S. Miller, 2014), few studies have studied the relation between ethics and the use of big data in practice. This research hypothesizes that security, privacy and ethics concerns do not play a role in the use of big data.

The factor **trading partner readiness** is about following partners like suppliers in the adoption of big data to maintain the ability for external collaborations (Sun et al., 2018). Within the literature this has been defined as the coercive pressure exerted by customers and suppliers to adopt a certain technology or system (Khalifa & Davison, 2006). In other words, the IT that partners use can influence the choice of IT for the firm (Fink, 1998). For example, a study found that the willingness to use e-commerce by SMEs was driven by the internet usage of its customers (Awa, Ukoha, & Emecheta, 2012). The expectation is that if the partners of a firm utilize big data, the SME will also use big data.

The factor **regulatory environment** is about the support government agencies offer to adopt big data (Sun et al., 2018). Policies that governments impose should be consistent and avoid duplication of effort to ensure effectiveness (Rashid & Al-Qirim, 2001). However, it is not just grants to stimulate innovation that comprise the regulatory environment (Fink, 1998). The regulatory environment should also aim to protect the organizational data with legislation (Oliveira, Thomas, & Espadanal, 2014). These normative pressures can guide the adoption of big data (Agrawal, 2015).

For example, in May 2018 the General Data Protection Regulation (GDPR) was implemented in Europe (IDC & The Lisbon Council, 2018b). It is supposed to tackle concerns regarding security and privacy of data. However, some argue that this new regulation is too strict for smaller business (The Economist, 2018). Therefore, the expected challenge is that regulations have a negative impact on the use of big data by SMEs.

New technologies are associated with the factor **uncertainty / risk concern**. These can arise as a result of unexpected consequences, expectations of profitability and the security concerns of big data (Sun et al., 2018). Unexpected consequences can lead to uncertain results, which translates to a certain amount of risk (Alshamaila et al., 2013). Furthermore, the steeper the learning curve of an innovation, the higher the risk is (Awa et al., 2012). As big data requires new skills, it can be inferred that it poses a high risk for firms.

Similarly, the factor **institutional based trust** is about the perceived safety of using big data (Sun et al., 2018). In this sense, it is the inverse of the **uncertainty / risk concern** factor. Trust in partners (Al-Qirim et al., 2017; Sila, 2013) and the reliability and safety of the platform that is utilized all play a role. In other words, it is about risks versus perceived safety and reliability. Therefore, these factors are merged together in the conceptual framework.

Not only trading partners can exert pressure on a firm. The factor **competitive pressure** describes the pressure from competitors on the firm to adopt new innovations like big data (Sun et al., 2018). It can force a firm to copy competitors in order to stay relevant (Khalifa & Davison, 2006). If competitors make use of certain IT (Fink, 1998), it can create a sense of urgency within the firm to adopt the same IT (Caesarius & Hohenthal, 2018). However, it is not just pressure from competitors.

The factor **market turbulence** is about the changing preferences of clients (Sun et al., 2018). SMEs need to be outward facing in order to listen to the wishes of clients and in order to formulate a vision (Ferneley & Bell, 2006). However, these pressures are exerted on the firm from the clients and via suppliers and competitors of the firm (Antlová, 2009). Big data offers the possibility to quickly adjust to this turbulence and personalize to the customer's need (Verma & Sekhar, 2017). While the initial pressure might come from customers, partners or competitors, **market turbulence** and **competitive pressure** are combined in a new factor called **market pressure**. While this research is interested in the origin for the pressure, it is better for the conceptual framework to simplify where possible. The expected effect is that the market pressure has positive influence on the use of big data by SMEs.

The final factor is **Information Systems fashion**. This is about the source of information for the acquisition of big data (Sun et al., 2018). If **market pressure** is about who exerted the pressure, then **IS fashion** is about where the information about big data came from. For example, due to the lack of business cases to follow as an example, firms are unable to determine how valuable big data is for their situation (Coleman et al., 2016). A lack of communication channels between vendors and clients will challenge SMEs and have a negative impact on big data adoption and use.

4.2 Summary

The factors that were previously discussed are combined into the conceptual framework below. Table 2 summarizes the factors in keywords and the expected effects. This way, the conceptual framework answers research question two: *what are the potential challenges of big data adoption?* The conceptual framework incorporates twelve factors, seven of which are expected to be a challenge to overcome in adopting big data: human resources, management support, technology resources, decision-making culture, regulatory environment, uncertainty / risk concern and IS fashion. Four factors are expected to positively affect the adoption of big data: change efficacy, IS strategy orientation, trading partner readiness and market pressure. One factor is expected to have no influence on big data adoption, which is concerns regarding security, privacy and ethics.

In the following chapters the conceptual framework will be applied to study the use of big data and challenges of adopting big data in SMEs, which is the focus of research question three and four respectively. Research question three is important to set the context, as a recurring theme is the question whether big data is unique and truly revolutionary for companies and their creation of competitive advantages. Looking at the conceptual framework in Table 2 does indicate that the factors influencing the big data use in firms are not unique to big data, but could apply to IT in general. Empirical study is necessary to investigate whether these factors might contain a different meaning in the context of big data use. Therefore, it is essential to establish whether this context is truly big data according to the definition of this research. The next chapter will discuss the methods for this investigation.

Table 2 - Conceptual framework

Nr	Category	Factor	Expected effects
1	OC	Human resources	Insufficient skills hinders use of big data
2	OC	Management support	Lack of support from management hinders use of big data
3	OC	Technology resources	Insufficient quality of existing infrastructure hinders use of big data
4	OC	Decision-making culture	A risk averse owner hinders use of big data
5	OC	Change efficacy	Flexible SMEs are likelier to use big data
6	OC	IS strategy orientation	Exploring options positively influences big data use
7	EC	Security, privacy and ethics concerns	Concerns on these issues do not influence big data use
8	EC	Trading partner readiness	Partners using big data positively influences big data use
9	EC	Regulatory environment	Regulations negatively influence big data use
10	EC	Uncertainty / risk concern	High uncertainty about big data (platform) reduces the use of big data
11	EC	Market pressure	Market pressure has positive influence on big data use
12	EC	IS fashion	Lack of communication channels will impact use of big data negatively

Chapter 5 – Methods

This chapter presents the methods used for the empirical investigation to answer research questions two and four. The findings on the second research question are presented in the next chapter: *why is big data a unique resource and how does it differ from traditional IT in Dutch consultancy SMEs?* The findings on the fourth research question are presented in chapter seven: *what factors influence the adoption of big data by Dutch consultancy SMEs?* This chapter will start with a justification of the research strategy, followed by the research design and data collection. The section after that discusses the method of data analysis and the chapter concludes with a discussion on the research quality and limitations of the research methods.

5.1 Research strategy justification

There are three steps in answering research question four. The first step was taken in the previous section, by constructing a conceptual model based on the literature review. The second step is developing an interview protocol based on the conceptual model. The third step is using this interview protocol for asking questions about the factors of interest. This final step is also used to answer research question three. The interviews were held with consultancy firms with 250 employees or less, as they are the focus of this thesis. During this process the interview protocol was continuously evaluated to ensure clarity for the interviewees.

There are several reasons for this research strategy. First, a qualitative research approach is associated with interpretivism (Bryman, 2012). This epistemological stance stresses that the social world can be studied by examining the interpretation of participants of the social world (Creswell, 2014). This requires a different approach, as this stance assumes interpretations cannot be quantified. Therefore, the focus is on understanding relations in terms of words (Bryman, 2012). As this research is interested in *why* firms are lagging behind in big data use, this approach is very well suited due to its focus on explaining a phenomenon (Bryman, 2012).

This leads to the second reason, which is the exploratory nature of this study. While there is general literature on the adoption, use and effects of big data, little is known about how big data is used in practice within SMEs and the specific motivations and challenges these firms face in reality (Coleman et al., 2016). A qualitative approach is well suited to study the perception of the problem SMEs face (Bryman, 2012). It is paramount to better understand the problem first, as solutions can only be formulated after these steps have been taken.

The third reason for the qualitative approach is the complexity of the problem (Bryman, 2012). While qualitative research is often associated with inductive reasoning and theory building, they are not intrinsically linked (Bryman, 2012). For example, quantitative interviews focus on the concerns of the interviewer and leave little room for deviation from the interview protocol. In contrast, qualitative interviews are about the concerns of the interviewee (Bryman, 2012). The complexity of the problem – as seen in the number of factors in the conceptual framework – requires a flexible approach like semi-structured interviews. This enables the investigation to ascertain the importance of factors that emerged from the literature, while leaving room for issues that might emerge from the interviewees during the processes. In this way, the complexity of the situation is made more manageable.

The final reason is related to that of the complexity and is about the context of the problem. There can be many different reasons and interpretations of the present issues, which are dependent on the context (Creswell, 2014). A qualitative approach is well suited to capture the meaning of data in its original context (Bryman, 2012). This ensures rich data that captures meaning as it is interpreted in the real world, in contrast to an artificial setting (Bryman, 2012).

Of course, qualitative research does have its drawbacks and critiques (Bryman, 2012). These will be point out here and discussed in more detail in the final section of this chapter. First of all, qualitative research risks being too subjective by allowing the researcher to deem what is important to focus on and what is not (Bryman, 2012). In contrast, quantitative research formulates focus points more explicitly with the use of existing literature (Creswell, 2014). To counteract this bias, this study started with explicitly stating the focus points in the existing literature. This increases the transparency for choosing certain factors over others.

A second critique is the difficulty to replicate qualitative research (Bryman, 2012). The evolving of interview protocols, the characteristics of interviewer and interviewee and their interaction are difficult to replicate for an outsider (Freese & Peterson, 2017). For example, a researcher might prioritize aspects during an interview that they empathize the most with (Bryman, 2012). However, some argue that replication is not as important for qualitative research due to its interpretive nature (Freese & Peterson, 2017).

Instead, they propose a different term: *dependability* (Bryman, 2012). The dependability is increased if records are kept of all stages of a research. For example, interview notes, formulation of research question and how interviews were coded. This enables outsiders to 'audit' a research and determine if the conclusions are logical (Bryman, 2012; Freese & Peterson, 2017). This also addresses the lack of transparency associated with qualitative research. While this is aimed more at how participants were selected and the research carried out, it is an important issue to discuss (Bryman, 2012) .

A final, major critique of qualitative research is the problem of generalization (Bryman, 2012). Critics argue that results of interviews with a small number of individuals in a particular context are impossible to generalize to other contexts (Bryman, 2012). In short, qualitative research is focused on acquiring depth and not breadth in data (Bryman, 2012). However, while qualitative research might not utilize the same statistical sampling methods as quantitative research, it is still possible to engage in *moderatum* generalizations (Bryman, 2012). This is only possible when the results of qualitative research are thick descriptions. This enables others to determine whether certain findings could be transferred to another context (Bryman, 2012).

In summary, qualitative research incurs critiques relating to subjectivity, lack of transparency and difficulty to replicate and generalize. Some critiques can be counteracted and others are more fundamental consequences of qualitative research. Nevertheless, this approach is well suited for the exploratory nature of this research into a complex and context dependent issue. The results cannot be generalized to an entire population, but does offer indicative insights and therefore adds understanding to the overall problem.

5.2 Research design & data collection

The previous section discussed the qualitative orientation of this study. However, the distinction between qualitative and quantitative does not indicate the design of the research and the methods employed to collect data. Therefore, this section will begin by discussing the chosen research design followed by a discussion of the methods employed.

A research design is a framework for collecting and analyzing data (Bryman, 2012) or in other words, the process in which the research is carried out (Creswell, 2014). Research designs differ by prioritizing certain dimensions of a research process. As a result, designs differ in the expression of causal relationships between variables, the ability to generalize findings to a larger population, having a temporal component in the research and how behavior is understood (Bryman, 2012). A research design is therefore a reflection of the importance attributed to each of these dimensions by the researcher. In turn, a research design has influence on the reliability, validity and ability to replicate a research (Bryman, 2012).

In this thesis, a cross-sectional research design is employed. A study with this design collects data on multiple cases at a single point in time and uses multiple variables (Bryman, 2012). It is associated with quantitative data, but can also be employed within a qualitative research strategy (Bryman, 2012). The data is analyzed by looking for patterns of association between variables. However, due to the lack of a temporal aspect, it is much more difficult to credibly infer causal relationships (Bryman, 2012). Instead, the goal is to determine whether there is a relation between variables.

In this exploratory research a conceptual framework is used to study possible influences on the use of big data by SMEs. This approach does not result in the determination of causal relationships between certain factors and big data use, but it is able to index important influences on big data use. The result of using multiple cases is that more variation is established in all the variables (Bryman, 2012). In turn, finer distinctions between cases can be made and saturation of the findings is likelier to occur with a larger sample size (Bryman, 2012).

The cross-sectional research design is well suited, due to the complexity of the issue. A case study might yield a richer description, but could completely ignore certain factors that are important in other cases (Bryman, 2012). A longitudinal study is more interested in the change over time (Creswell, 2014), while this study is interested with establishing important factors first. Similarly, comparative research would study the similarities and differences between cases (Bryman, 2012), but is not as well suited for determining why or how factors influence big data use.

Now that the research strategy and design are discussed, it is time to turn to the research methods. The research method is the instrument that is used for the collection of data (Creswell, 2014). In this thesis, semi-structured interviews were held to collect data. This form of interview is between structured or questionnaire-style interviews and unstructured or open interviews (Bryman, 2012). The advantages of this method are the structure it provides for the researcher and the interview process, while leaving room for the participant to add insights outside of the predetermined structure (Creswell, 2014). This ensures that important topics will be covered, but leaves the opportunity to change the order of questions and follow up on things said by the participant (Bryman, 2012). The flexibility and ability to explore topics of interest of the participant make semi-structured interviews the method of choice for this thesis.

The interviews were held retrospectively, meaning that participants looked back on how they experienced past events (Bryman, 2012). This does introduce the problem of recall bias, in which it is difficult to accurately remember past events. However, decreasing the time between the interview and the object of study can reduce the bias (Bryman, 2012). To reduce the bias in this thesis, participants were asked how long ago they started using big data. The shorter the time period, the less likely a recall bias is to occur.

There were several criteria that had to be met to participate in this research. This ensures a similarity between the participants and a sample that can be used to answer the research question. The first criterion is that the firm, the unit of analysis, has to have 250 employees or less. This ensures it falls in the SME category. The second criterion is that the firm has to operate in the consultancy or IT services sector. The third criterion is that the firm has to (claim to) use big data in products or services it offers.

Research questions in qualitative research should indicate what the unit is that needs to be sampled (Bryman, 2012). In this research, SMEs in the consultancy sector are the unit of analysis. To get a sample of these firms, purposive sampling is used (Bryman, 2012). With this method, the researcher applies a non-probability form of sampling. Instead of randomly selecting participants, interviewees are chosen that are relevant to the research. These can vary in terms of characteristics like firm size and age, but do have to be useful in answering the research question. The downside of this approach is the inability to generalize to a larger population (Bryman, 2012). Nevertheless, for exploratory research it is a useful method.

The search for participants started on Google. With search terms as *data science firm*, *big data firm* and *big data consultancy* an initial list was made of possible SMEs. Representatives of firms were approached for an interview. Interviews were held with either management of big data related departments or with data scientists working with big data. These offer different but valuable perspectives. Once someone agreed to an interview, a date, time and method of interview were discussed. Interviews were held either face-to-face or via telephone.

The participation of the interviewees is anonymous and they were only recorded if permission was given. The purpose of the recording was to aid in the transcription process. Notes were also taken during the interview and if there was any uncertainty during transcription, the interviewee was contacted. Interviews ranged between 30 minutes to over an hour, depending on the individual, available time and whether it was face-to-face or via telephone. This influenced time-management and the depth of the interviews. The interview protocol provided guidance and a time-management tool by covering the necessary questions. In addition, after some interviews it was possible to estimate how much time the remaining questions would take. Interviews were held in Dutch, unless the interviewee preferred English.

The interview guide itself is based on the conceptual framework and can be found in Appendix B. It starts with introductory questions about the role of the participant in the firm and about what the firm does. It then moves to questions relating to the specific factors of the conceptual framework and whether they have any influence according to the interviewee. While the questions are numbered, the order in which they are asked is flexible and dependent on the previous responses of the participant. Furthermore, a question could be left out if a participant made remarks earlier about a certain topic.

5.3 Data analysis

Analysis of qualitative data consists of preparing, analyzing and understanding the data (Creswell, 2014). Therefore, the recordings were transcribed as soon as possible. The transcripts were divided in snippets of text and placed under their corresponding questions in the interview protocol. This structured the transcripts and made the analysis easier. The transcripts were coded with the use of NVivo coding software. The transcripts were thoroughly read and the factors influencing the adoption or use of big data were marked. Coding the transcripts is a way of structuring information (Creswell, 2014). It also is a way of better understanding the data and dive deeper in hidden layers of meaning. This leads to an interpretation of the overall data (Bryman, 2012).

The coding process is done as soon as possible but only after reading the transcript completely for the first time. The whole process followed the considerations of Bryman (2012). For example, multiple codes might apply to the same text segments. However, it is important to review codes and keep in mind that it is only a step in the overall analysis. The result is more organized data that can be interpreted by the researcher (Bryman, 2012).

The codes are based on the conceptual framework and shown in Table 3. The codes of the IC category were also added, as was an extra code to mark text describing the definition of big data or which technologies are used. These additions were necessary to answer the third research question and determine whether the IC were of any influence on big data adoption. Furthermore, if factors emerged from the interviews that did not fit in any of the categories, it was coded as 'other'. The codes themselves do not yet clarify whether a factor was of positive or negative influence, this interpretation was done after all the transcripts were coded.

During the interviews, the coding protocol was continuously evaluated. This ensured that factors that emerged from interviews, but did not occur in the framework, were incorporated for future interviews. The process of coding is illustrated in Appendix C with some examples.

Table 3 - Coding protocol

No.	Factor	Code
1	Human resources	HR
2	Management support	MS
3	Technology resources	TRes
4	Decision-making culture	DeCu
5	Change efficacy	CE
6	IS strategy orientation	ISO
7	Security, privacy and ethics concerns	SPEC
8	Trading partner readiness	TPR
9	Regulatory environment	RE
10	Uncertainty / risk concern	URC
11	Market pressure	MPR
12	IS fashion	ISF
13	Relative advantage	RA
14	Compatibility	Cpat
15	Complexity	Cplex
16	Trialability	Trial
17	Observability	Obs
18	Cost of adoption	CoA
19	Big data definition	BDD
20	Other	Oth

5.4 Research quality

The validity and reliability of this research is increased in various ways. This is necessary, as this research is exploratory in nature and therefore it is more difficult to reliably confirm its findings. In qualitative research, validity refers to the checking of the accuracy of the findings by using certain procedures (Creswell, 2014). Several types of validity exist and the three most important are internal-, external- and measurement validity (Bryman, 2012). Generally, internal validity is about causality and the extent conclusions can be drawn based on the research (Bryman, 2012). External validity is about the issue of generalization and whether the findings of a research can be generalized to other contexts (Creswell, 2014). Measurement validity is less of an issue for qualitative research, as it is concerned with finding valid measurements or indicators to accurately reflect concepts (Bryman, 2012).

Several different strategies can be used to increase the validity of a research. These include the triangulation of data sources, member-checking, bias clarification and peer debriefing (Creswell, 2014). Triangulation is about constructing a theme from different data sources. A literature review consisting of academic and grey literature combined with interviews with people from different firms, levels and backgrounds results in an abundance of data. If this data from multiple sources converges on a theme, then you can a higher research validity (Creswell, 2014). The articles and methods used for triangulation can be found in Appendix D.

The interviews of this research are held with as many SMEs as possible. Within these firms, either a management-level employee or someone involved in data science, data engineering or business intelligence with practical knowledge about big data are interviewed. This ensures both a larger sample size and a variety of perspectives, increasing the richness of the data and providing more insights. In very small firms, managers are often also the person in charge of actually working with big data.

Member checking is about determining the accuracy of the findings by allowing participants to check a draft and provide them with the opportunity to comment (Creswell, 2014). If they agree with the findings, this increases the validity. If they do not, the researcher should discuss this with them to find out why (Bryman, 2012). In this research, participants were sent a follow-up e-mail after the interview with a summary of the findings so far and a request to comment if they agreed.

Bias clarification is about honesty and self-reflection of the researcher (Creswell, 2014). The researcher should be open and clear about possible influences. These might influence the interpretation of the findings and could be shaped by a variety of factors including culture, socioeconomic origin and whether it was a paid research (Creswell, 2014). For example, this thesis was written as part of an internship at Technopolis BV, but does not represent any value for the organization. Rather, the desk research I did during the internship helped me find a literature gap and formulate a research proposal.

Peer debriefing enhances the accuracy of the research by finding a person that is willing to critically assess the research and ask questions about the methods and findings (Creswell, 2014). This research has used the network and expertise provided by the Technopolis Group, an expert workshop and peer-reviews to strengthen its validity.

The Technopolis Group has an extensive network and employs experts in the domains of science, technology and innovation. More specifically, I worked with and got feedback from people that have specialized their knowledge in digital technologies like big data, SMEs, entrepreneurship and skill development. These people are knowledgeable and up-to-date on recent developments and feedback acquired from talking with these people was used in guiding and validating this research and its findings. In particular their experience with interviews was utilized in designing the interview protocol.

As part of the overarching project an expert workshop was organized on the 6th of June 2018 in Brussels. The topic of this workshop was how skill development could facilitate the uptake of technologies like big data, Internet of Things (IoT) and cybersecurity. It attracted experts from all over Europe with different backgrounds. In total, 21 experts participated to discuss the preliminary findings of the overarching project. Due to the different expertise and backgrounds of the participants, the subject also covered other topics than skill developments. Therefore, this expert workshop was also used in validating the findings.

A peer-review was held to strengthen the validity of this research (Creswell, 2014). Discussing the preliminary findings with fellow graduate students provided valuable and constructive feedback. This was used to fine-tune the research and provide input for the discussion and conclusion.

Qualitative reliability is concerned with employing a consistent research approach across different researchers and projects (Creswell, 2014). It is about the question whether the results of a research are repeatable (Bryman, 2012). The reliability of a research can be increased by carefully documenting all steps taken to get the results (Creswell, 2014). It also ensures accountability and enables others to replicate the study to see if the results are the same. In this qualitative study reliability is ensured by providing a detailed method section for both the literature review and interviews and making use of peer-reviews and code checking during the analysis. The latter method ensures snippets of text are coded the same, regardless of the researcher. A fellow graduate student was asked to check the codes in the anonymous transcripts to see if the student would get to the same results.

Chapter 6 – Big data in Dutch SMEs

This chapter presents the first results of the empirical investigation as described in the previous chapter. The focus of this chapter will be on answering the empirical part of the second research question: *why is big data a unique resource and how does it differ from traditional IT in Dutch consultancy SMEs?* It is important to understand the use and application of big data, as this context determines whether the conceptual framework is unique for big data or not. This contributes to putting the findings of the fourth research question in the appropriate context.

The chapter begins by discussing the setting in which the investigation was performed. This is followed by a discussion of the big data projects of the interviewees, which illustrates how big data is used in practice. Based on these results and the definition of big data as presented in chapter 2 an overview can be made. This illustrates the contrast between theory and practice when it comes to big data. It also highlights the differences between the firms and contributes to understanding whether big data is truly unique in the context of this research.

6.1 The Dutch context

The interviews have been undertaken between July 2018 and September 2018 with representatives from several small and medium enterprises in The Netherlands. The choice for the Netherlands as a geographic delineation is chosen because it emerged from the overarching project that it has a relatively well-developed big data ecosystem. The Netherlands has SMEs involved in data science and support institutions like the Dutch Chamber of Commerce promoting the use of digital technologies (Van der Veen, Van den Born, Smetsers, & Bosma, 2017).

The second reason for The Netherlands is the high percentage of big data already being used in firms of all sizes (Centraal Bureau voor de Statistiek, 2017). In fact, the country has the highest percentage of firms using data within the EU (IDC & The Lisbon Council, 2018a). For example, 20% of micro-sized firms have used big data analytics themselves (Centraal Bureau voor de Statistiek, 2017). This percentage is high enough to allow for interesting cases, but still leaves room for improvement and learning opportunities.

This leads to the third reason. The Netherlands is not in the top-5 of European countries in which the data market is concentrated (Davies, 2016). Even though the percentage of firms using or having used big data is high, the country is lagging behind Germany, France, Italy, Spain and the United Kingdom (IDC & The Lisbon Council, 2018a). Thus, by understanding the barriers that SMEs in The Netherlands face in the deployment of big data, this research can contribute to the overall competitiveness and position of the country in Europe and worldwide.

The type of SME approached for the interviews are firms active in the IT services and consultancy sectors. Examples about how big data could transform large manufacturing firms are plentiful (Lee, Kao, & Yang, 2014; Manyika et al., 2011; Mazzei & Noble, 2017), but less is known about small service firms. The choice for IT- and consultancy-service firms was made due to their business model. These firms depend on the knowledge and skill of their personnel (Obeidat et al., 2016). As the data ladder of Roden et al. (2017) in Figure 2 depicts, data leads via information to knowledge. Consequently, these firms are great candidates for using big data to complement their current business model (Christensen et al., 2013). These firms might have applied big data for themselves or for their clients. Interviewing them will shed light on the factors that they encountered in adopting big data. In turn, other firms could use this knowledge to increase the likelihood of successfully adopting big data.

Most of the frontrunners in terms of big data application are in the IT sector (Van der Veen et al., 2017). Firms outside the IT sector could have hired a third party to provide big data as a service (Ardagna et al., 2016) or might use the technology to supplement their core business. The latter is likely in sectors that are relatively well-developed in data use in the Netherlands, like industry, finance and logistics (Van der Veen et al., 2017). Both IT and consultancy firms were approached to participate in an interview.

Table 4 indicates how large the IT and consulting sectors are in the Netherlands. A total of over 1.7 million firms are operating in the Netherlands, of which 67 thousand operate in IT and almost 120 thousand in management consulting (Centraal Bureau voor de Statistiek, 2018). Together, these sectors account for almost 11% of all firms in the Netherlands. However, the majority of these firms consist of four employees (abbreviated 'Empl' in Table 4) or less. In IT more than 92% of firms and in management consulting 98% belong to this small category. This thesis focuses on the organizational resources and dynamic capabilities. A minimum of five employees was therefore applied in the selection of interviewees. This left a total of 4870 firms in the IT sector and 2475 in the consulting sector with more than five employees.

Table 4 - Number of firms in the Netherlands per sector and number of employees (CBS, 2018)

Sector	**Total no. of firms**	**5-10 Empl**	**10-20 Empl**	**20-50 Empl**	**50-100 Empl**	**100+ Empl**
All firms	1.726.190	62.940	33.055	20.100	7.095	7.875
IT	67.380	2.005	1.340	990	300	235
Consulting	119.935	1.325	640	340	110	60

The starting point of the sample selection was to search for Dutch data science firms on Google. These firms provide data science services to other firms. They might have evolved to incorporate big data technologies as well and could therefore provide insight in how the decision was made and what influenced it. The search for big data firms in the Netherlands resulted in a list of 30 firms. Representatives of these firms were approached with either an email, LinkedIn message or telephone call. Personal contact was deemed best to persuade people to participate, so individual employees of a firm were approached.

Individuals were selected on their role within the firm. As the research focuses on organizational resources and dynamic capabilities, individuals in management roles were initially approached. If they lacked the time or knowledge, people in a data science role within the organization were approached. A snowballing method (Bryman, 2012) was also used to get in contact with more companies from the interviewee's network, which resulted in two more firms for the list.

In total, 14 firms agreed to an interview, ranging in size from 5 to 150 people (see Table 5). Four firms fall in the category 5-10 employees, two firms with 10 to 20 employees, five firms with 20 to 50 employees, two firms with 50 to 100 employees and one with more than 100 employees. The overall composition of the chosen sectors in the Netherlands indicates that this is a fair sample considering the number of firms per employee category. Due to the nature of the research it is not representative, but does enable the formulation of indicative results.

Table 5 - Firms in sample

No.	Industry	Year founded	No. of people	Interviewee's position
1	IT Consultancy	2015	28	Business Manager
2	IT Consultancy	2013	5	CEO
3	Management Consulting	1997	20	Senior Advisor & Advisor
4	Information Services	2000	61	Big Data Engineer
5	IT	2017	8	Founder & CTO
6	Management Consulting	2004	65	Senior Data Scientist
7	IT	2008	7	Data Scientist
8	Management Consulting	2001	150	Director
9	IT	2014	14	Director
10	IT Consultancy	2012	25	Team Lead Data Science
11	IT Consultancy	2011	26	General Manager
12	IT Consultancy	2011	25	Software Engineer
13	Marketing Consultancy	2009	5	Managing Director
14	IT	2003	40	CEO

6.2 Big data in Dutch SMEs

There is certain variety in the firms and their application of big data. Some provide coding services, while others implement specific big data solutions for their clients. These include algorithms, BDAaaS or evidence-based advice. The wide variety of interpretations of big data brings to light a discrepancy between theory and practice. As discussed in chapter 2, the term big data originated in the late '90s and the definition remains vague and fuzzy. In this thesis it is defined as: "*high volume, high velocity or high variety information assets*". This still leaves room for interpretation on what 'high' means, however, which is dependent on time and context due to constant technological progression. This section will briefly discuss the characteristics of big data the interviewees ('Int.') utilized and provide an overview for each interviewee.

Figure 5 shows the results of the coding process for this part of the research. The codes that were used related to the six IC characteristics as formulated in chapter three and an extra code to capture remarks on the definition or use of big data within the firm. This code occurred the most, as it is the only one across all interviews and with a total number of 85 references. This is a logical consequence of the interview protocol, as it starts with a question about big data within the firm. The remaining factors were not explicitly asked, but only emerged due to the semi-structured nature of the interviews.

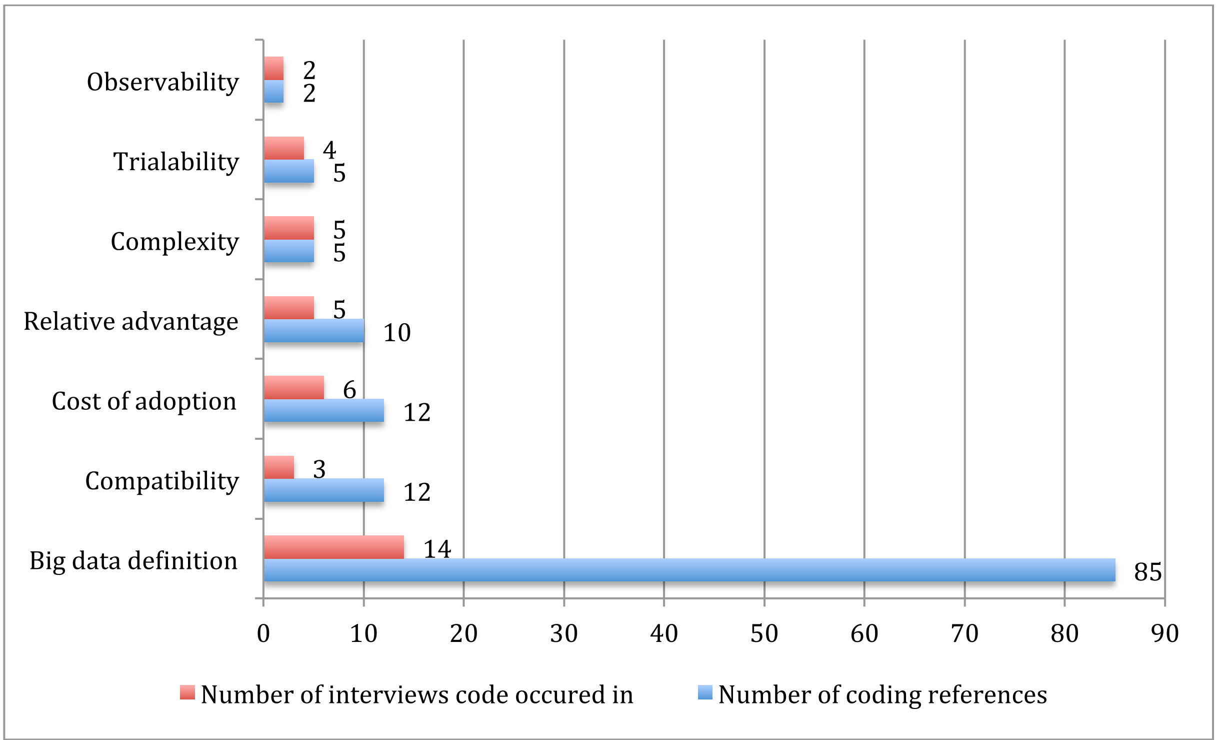

Figure 5 - Number of coding references for IC

Starting with the **big data definition**, in which every interviewee questioned whether size was the most important aspect. Or, as one interviewee put it: "*I can tell you how much data we use and then you can decide whether it is big*" (Int. 7). There was no consensus on a threshold in terms of volume that would turn data into 'big' data. One interviewee stated: "*For some people big data is something really big in size, that does not even fit into one disk. So terabytes of data (...) but in practice, they mean something that does not fit into Excel* " (Int. 2). This indicates that new software or other tools are needed to process these datasets, which is sometimes also used to define big data (Cavanillas et al., 2016).

This definition of big data states that traditional technologies are unable to cope with the challenges presented by big data. The technologies utilized by the interviewees paint a picture that either these traditional technologies have caught up or that the data is not truly 'big'. Four firms make use of the software program R for analysis of data (Int. 2, 6, 7, 9), which in itself is not developed specifically for big data.

The database management software Hadoop is specific to big data and used by interviewees 4 and 8, while interviewee 7 is the only one to use MapReduce. Other technologies and programming languages mentioned are Google Cloud, Python, SPSS, PowerBI, Java, machine learning and algorithms in general. Not all of these technologies could be considered solely with big data in mind. For example: "*You can do some advanced stuff with only a thousand records and still get some cool insights*" (Int. 9). Therefore, it is not sufficient to only look at the technologies used by the firms to determine whether they use big data.

However, two interviewees actually did link the technologies and the definition of big data together. They stated that: "*it is something that you cannot process with traditional software like Excel or even small database software*" (Int. 2) and "*You should actually call it difficult data (...) big data is where the traditional databases, warehousing, business intelligence and other systems are insufficient*" (Int. 4). These interviewees represented consultancy firms that have programming and Customer Relationship Management (CRM) as their core business.

For others, the type of data was more important than the actual size. One firm is building an online database of photos, uploaded by its customers. With this database it is developing an algorithm that could select the photos best suited for printing (Int. 5). Three other firms consider data to be big if it is unstructured or composed of multiple data sources (Int. 8, 10, 11). The use of the data is important as well: "*It is not about the size of the data, but about its use. (...) What are the use cases, can you build predictive models?*" (Int. 1). This refers to an often-mentioned additional V: value (Cavanillas et al., 2016).

One interviewee advocated for changing the term to 'difficult data' as big is dependent on time and context: "*It is an ongoing process, what was big fifteen years ago, you can do on your phone now*" (Int. 4). Another said: "*Big data has been around for more than ten years now*" (Int. 2). Yet another responded with: "*I do not really care what big data is, you can either scale your system or speed up your algorithms, there is always a solution*" (Int. 7). This would almost make it seem that big data is a problem of the past. However, perception seems to be very important: "*For us, 100 million records is already big, while others might not think so*" (Int. 3).

These findings indicate that big data is as fuzzy in reality as the theory indicates. Some firms utilize technologies specifically developed for big data, strengthening the argument that big data is different from traditional IT. However, the majority of firms use technologies not specific to big data. This, combined with the different perceptions of what big data is, requires a further look at the remaining IC codes and the context of the application of big data.

Of course, the IC of Rogers (2003) influence the adoption rate of an innovation. These might be considered challenges and are therefore discussed in more detail in the next chapter. Three factors deserve attention in this chapter however: complexity, compatibility and relative advantage. Big data is more complex than IT due to higher reliance on statistics, math and programming skills. Second, big data systems might not be compatible with other IT systems.

This issue occurs between general IT systems as well, but combined with the higher complexity it is a distinctive characteristic. The final argument for the uniqueness of big data is the relative advantage it offers compared to traditional IT solutions. With the right amount of data, an unprecedented amount of correlation and precision can be acquired (Cukier & Mayer-Schoenberger, 2013). This was simply not possible before its introduction. These arguments make a case for the unique setting of big data, but the next sections and chapters have to determine whether it truly is.

Table 6 on the next page shows an overview of the projects that each of the interviewees considered big data. It also indicates how big data was applied, which will be discussed in the next section. Lastly, it notes which V it incorporates of the 3V model and whether this qualifies as high. If it is, then it can be considered big data. Eleven firms mentioned volume as the defining V or in combination with another. Among the interviewed firms, nine can be considered to use big data. The remaining firms do not have data projects that qualify as big data within the given definition.

Table 6 - Big data use in firms

Firm no.	Description of big data	Which V's?	Big data?	Application
1	Not size, but the use case is important	Volume	Yes	Process optimization with multiple goals: customer experience, working more efficiently, saving costs
2	Something that does not fit in excel and cannot be processed on a single computer	Volume	No	Maintenance of technological infrastructure and (B)DAaaS
3	Data with half a million rows	Volume	No	Increasing validity of product (advice) with data
4	Big data is predicting, difficult data is about volume	Volume	Yes	Increasing validity of product (advice) with data
5	Hundreds of thousands of photos	Variety, Volume	Yes	Improving customer experience
6	Data on three hundred thousand people	Volume	No	Increasing validity of product (advice) with data
7	Five hundred million rows per day	Volume, Velocity	Yes	Increasing validity of product (advice) with data
8	Collection of structured and unstructured data	Variety	Yes	Increasing validity of product (advice) with data.
9	Using advanced statistical analyses	Volume	No	Process optimization with multiple goals: customer experience, working more efficiently, saving costs
10	Connecting different data sources	Variety	Yes	Extracting (hidden) insights from big data
11	Unstructured data with no use in mind that is transcending more than one application	Variety	Yes	Extracting (hidden) insights from big data
12	Too large to use on a single computer	Volume	No	Process optimization with multiple goals: customer experience, working more efficiently, saving costs
13	Reach of 1.7 million people	Volume	Yes	Targeted marketing
14	Connecting systems with multiple data sources to each other. Has data on 3.5 million people	Volume, Variety	Yes	Process optimization with multiple goals: customer experience, working more efficiently, saving costs

A graphical representation can be seen in Figure 6, which is based on the 3V model. The further outward a dot is, the more it leans towards 'high' in one of the characteristics. It is not an exact representation, but interpretations of the described projects and how they rank relative to each other. For example, interviewees 8, 10 and 11 utilize unstructured data, so they rank in the outer circle on the variety axis. Meanwhile, interviewees 5 and 14 described unstructured data that is also high in volume – millions of row or hundreds of thousands of photos. As a result, they are in the outer circle between volume and variety.

Interviewees 2, 3, 6, 9 and 12 described data sources with only hundreds of thousands of rows in an excel file or admitted that they do not consider the data they work with to be 'big'. They rank therefore in the inner circle, considering them lower in volume than interviewee 13, which uses databases with a few million rows. In turn, interviewees 1 and 4 have such high volume data sources that they do not fit in a single computer hard drive. Finally, interviewee 7 ranks high both in terms of volume and velocity, as their data consists of half a million rows and is generated each day.

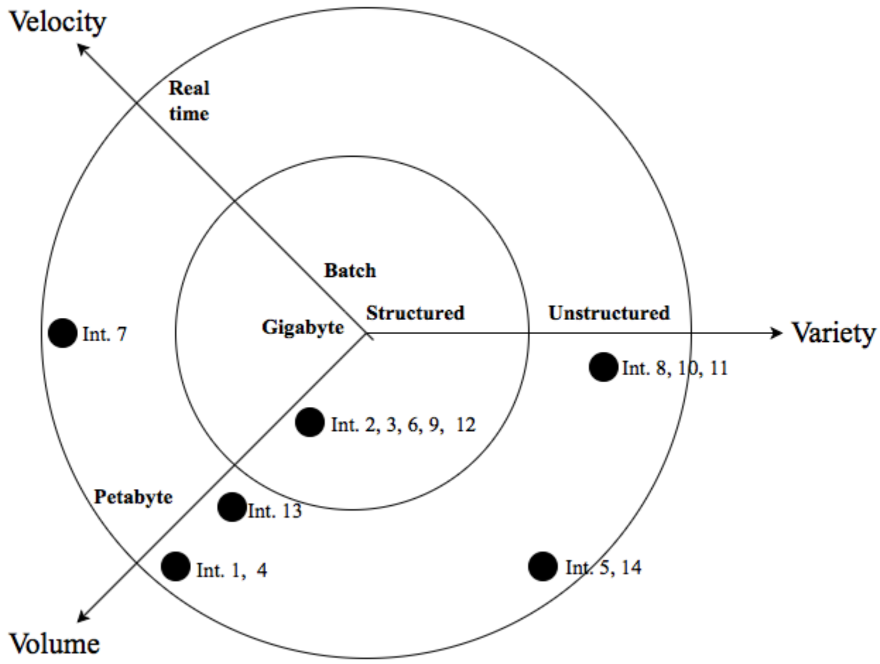

Figure 6 - Characteristics of each interviewee's big data project

The size is often still considered the primary characteristic of big data. No projects were described where all characteristics were high and only a few had overlap with multiple characteristics. For example: big data consisting of both high variety and high volume information assets. Also lacking was a project that combined unstructured data with a high velocity. An example could be the analysis of video or music files in or near real-time.

What was interesting to note, is that some interviewees considered technologies to be part of the definition of big data. As discussed in chapter 2, it is questionable to include the tools used in the definition. Nevertheless, the aspect of needing specific technologies to tackle big data's problems could indicate the need to expand the definition to include the complexity of big data.

6.3 Applications of big data

The previous section discussed whether the interviewed firms met the definition of big data. To understand if big data is unique and different from traditional IT, it is also necessary to look at its applications among the interviewed firms. This section will briefly discuss the goals of big data, to illustrate how it is used in practice and get an indication of its relationship with organizational resources and dynamic capabilities. This will give a better understanding of the overview presented in Table 6. Based on the interviews, big data is used for only a select number of reasons.

First, data is used to increase insight in certain topics, but the optimization of processes within the firm is the most cited. Four firms (Int. 1, 9, 12 and 14) used big data themselves or for their clients to analyze processes in the hope of increasing efficiency in terms of time or costs. These were specifically focused on optimizing the current processes, but not necessarily on the development of new value propositions. The fact that these firms have multiple goals could indicate the use of big data. As one interviewee said: "*big data has to transcend one application*" (Int. 11). He argued that as long as data has a single, predefined application you could not consider it big data.

When big data is used to optimize processes, it indicates the presence of dynamic capabilities (Shan et al., 2018). As discussed in chapter two, dynamic capabilities can be fostered by learning and result in a change of routines. BDA increases understanding of a process and creates an opportunity for learning and improvement. If the process is actually changed or optimized, it indicates that dynamic capabilities have been created as a result of big data. It is not unique to big data however, as IT has also been linked to developing capabilities (Real, Leal, & Roldán, 2006; Tippins & Sohi, 2003).

Similar to the aforementioned firms, interviewees 3, 4, 6, 7 and 8 utilize big data to increase the validity of their products or services. This revolves mostly around monetizing data by incorporating it into products or services. In other words, data is the raw material imported by the firm, transformed in operations and delivered to the client. Depending on the firm, data is used in different ways. For consultancies, data has become part of their own business model, basing their advice on data and evidence. The idea is that the more data you have on a phenomenon, the easier it is to describe and understand it (McAfee & Brynjolfsson, 2012). The next step is to be able to predict what will happen next (Cukier & Mayer-Schoenberger, 2013). Together, this will increase the (idea of) validity of the advice. As one interviewee said: "*We do mention to our clients we are able to use big data in order to distinguish ourselves from the competition*" (Int. 3).

With this application, big data is used as an organizational resource. In the case of utilizing BDA as a service, big data is a resource that the firm is exploiting by selling it on the market. The question is whether this benefits or changes the routines of the client it is sold to. If it is, then the client clearly possesses the capability to change routines as a result of externally acquired advice. If routines do not change it can have several reasons, like lack of feasibility, or capabilities were simply not altered.

Only interviewees 10 and 11 indicated that big data was used without explicit applications in mind. This distinguishes them from the other firms that mentioned big data was used for multiple goals. Firms 10 and 11 are focused on extracting insights from the data their clients are not utilizing. Their clients have data they do not exploit and these IT consultancy firms are tasked with organizing the data and identifying opportunities or use cases. As interviewee 11 said: "*Clients come to us because they have a lot of data they do not use and ask us to find out whether there are any connections or implicit information they could use*". When utilizing big data in this way, the data comes in and is analyzed, but could impact the client of the firm in any of its routines. It all depends on the information the data holds.

The final three firms had three different goals in mind with big data. Interviewee 2 provided services related to the maintenance of technological infrastructures and BDAaaS. This firm does not utilize big data for itself, but provides a combination of organizing information at its clients and extracting insights from the available data. Interviewee 5 possesses big data in the form of photos. The amount and type of data make it difficult to work with, so this firm is developing its own machine-learning algorithm to increase the experience for its customers. This algorithm has the goal of distinguishing good photos from bad ones, helping its customers in the selection of photos and increasing the customer experience. While the data of the firm is still provided by its customers, the customer experience is mostly impacted in the later stages of the value chain. It provides a different value proposition that marketing & sales can take advantage of and it will also change service from an after-sale perspective to something that is offered during the production process.

Finally, only interviewee 13 has data with the single and explicit goal of targeted marketing. This enables the firm to understand their customers better and act accordingly. Data on customers is used to market similar products and services the customer could be interested in, based on a variety of known characteristics. The upselling and cross selling of goods and services this enables translates to increased revenue. With this method, big data is used to leverage other organizational resources.

6.4 Summary

This chapter started with a discussion of the Dutch context in which this research was performed. The Netherlands was chosen because of its big data ecosystem, high percentage of data use among firms and competitive economy. Consultancy and IT service firms were approached, as their business model is based on knowledge and skill, which big data could complement. In total, fourteen SMEs were interviewed in the IT and consulting sectors about their use of big data.

The next section focused on what the interviewees thought of big data and whether their use of data was actually 'big' according to the definition: *big data is high volume, high velocity or high variety information assets*. A closer look at the data used by the interviewees shows that none are 'high' in all of the characteristics of big data. Most firms consider big data only in terms of volume, with few branching out to different data formats. The speed in which the data is gathered rarely plays a role and most of the data can be processed with traditional database management systems. A few firms do use specialized software, but for specific goals a smaller database might yield equally insightful results as a big dataset. Nine firms qualified for the definition of big data.

This chapter answered the empirical part of research question two: *why is big data a unique resource and how does it differ from traditional IT in Dutch consultancy SMEs?* In summary, the big data in the interviewee sample is used mostly as an organizational resource. The promise of transforming to data-driven organizations does not seem to have reached these firms yet. Nevertheless, they are experimenting with extracting insights and hidden information, which could very well impact other areas and contribute to learning and capability development. For example, using big data for understanding and optimizing processes could lead to a change in routines. In turn, this indicates the development of dynamic capabilities. However, none of the interviewed firms have drastically altered decision-making and basing it on big data. Big data is currently complementing the decision-making process with evidence and data.

The evidence for big data being more important for building capabilities and knowledge than as a resource is inconclusive. One interviewee did mention that reflection of the used methods is important (Int. 6) and others experimented with big data, expanding their knowledge and capabilities on the subject (Int. 3 & 4). However, for the majority of interviewees, big data still represented a valuable resource rather than a mechanism for fostering capabilities. As a result, the claim that big data is different from information technologies preceding it is not backed by these findings.

The question is how typical these firms and their projects are for the Dutch context. All these firms were chosen because they considered themselves users of big data. In their use of data, they do belong to the majority of data-using firms in the Netherlands (IDC & The Lisbon Council, 2018a). However, statistics on the use of *big* data in the Netherlands are scarce. It is likely that these firms are using big data more than the average Dutch SME, especially considering the sectors in which they are active (Van der Veen et al., 2017). A poll by the Dutch chamber of commerce found that 13% of entrepreneurs is utilizing big data (Van der Veen et al., 2017) Furthermore, the definition also determines the cutoff point between data and big data. The definition in this thesis is broad and therefore inclusive.

Chapter 7 – Factors influencing big data adoption in Dutch SMEs

This chapter will present the main results of the empirical investigation as described in chapter five. In doing so, it will provide an answer to the fourth research question: *what factors influence the adoption of big data by Dutch consultancy SMEs?* Where the previous chapter was concerned in determining how unique the development of big data currently is, this chapter will ascertain what issues are perceived in the adoption of big data by Dutch SMEs. First, the results from the interviews are presented. The effects on the adoption and use of big data are discussed. Then, an interpretation is made of these findings in the context of the conceptual model. The final section will answer the fourth research question.

The conceptual framework as described in chapter four is used to see whether the expected effects are found in reality. In turn, this generates insights in the relationships as proposed in the research model in chapter two. For a better overview, these two are combined in the conceptual model as shown in Figure 7. This model incorporates the factors of the framework within the proposed research model of organizational resources and dynamic capabilities.

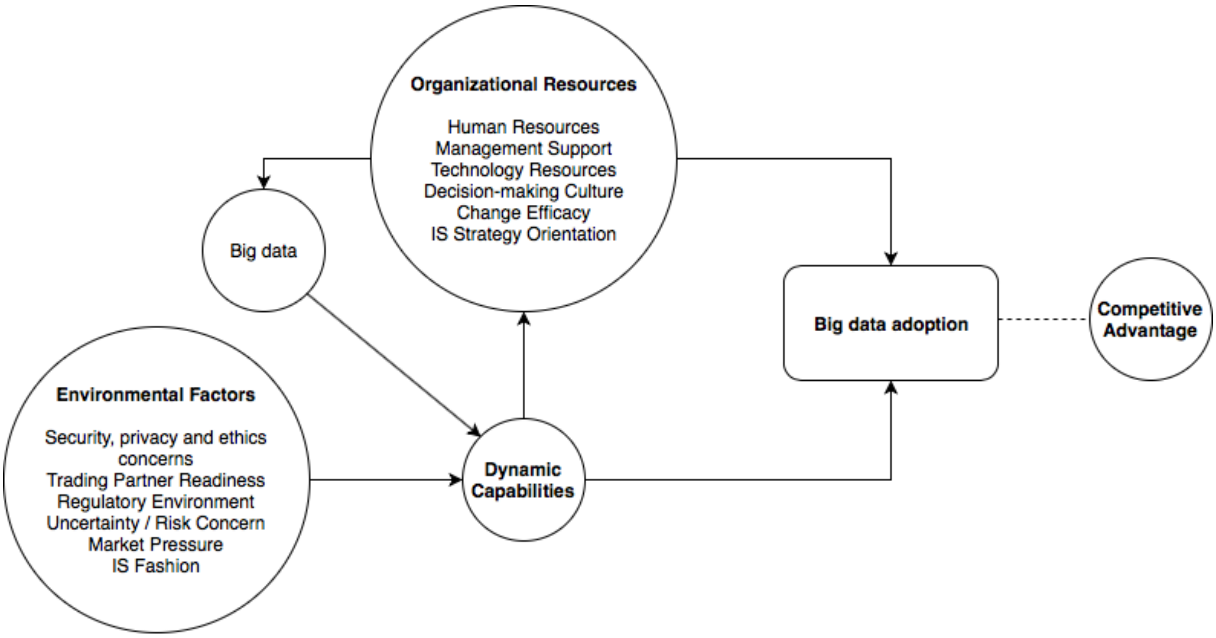

Figure 7 - Conceptual model

The coding protocol discussed in the methods chapter was applied to code and interpret the interviews. Figure 8 shows the total number of times each factor was marked in the transcripts. It also shows the number of interviews a factor occurred in. Together they provide an overview of the most covered factors. Due to the semi-structured nature of the interviews, the questions and how the interviewee interpreted these influence the number of references for certain factors. There was no explicit question about IS fashion for example, but there was a question on market pressure. The difference in occurrence of these factors can thus partly be explained by the interview protocol itself. Nevertheless, if the interviewee did not bring up a certain factor, it can also be assumed that it is of no (perceived) importance. The rest of this chapter will briefly discuss the individual factors and what interviewees said about them.

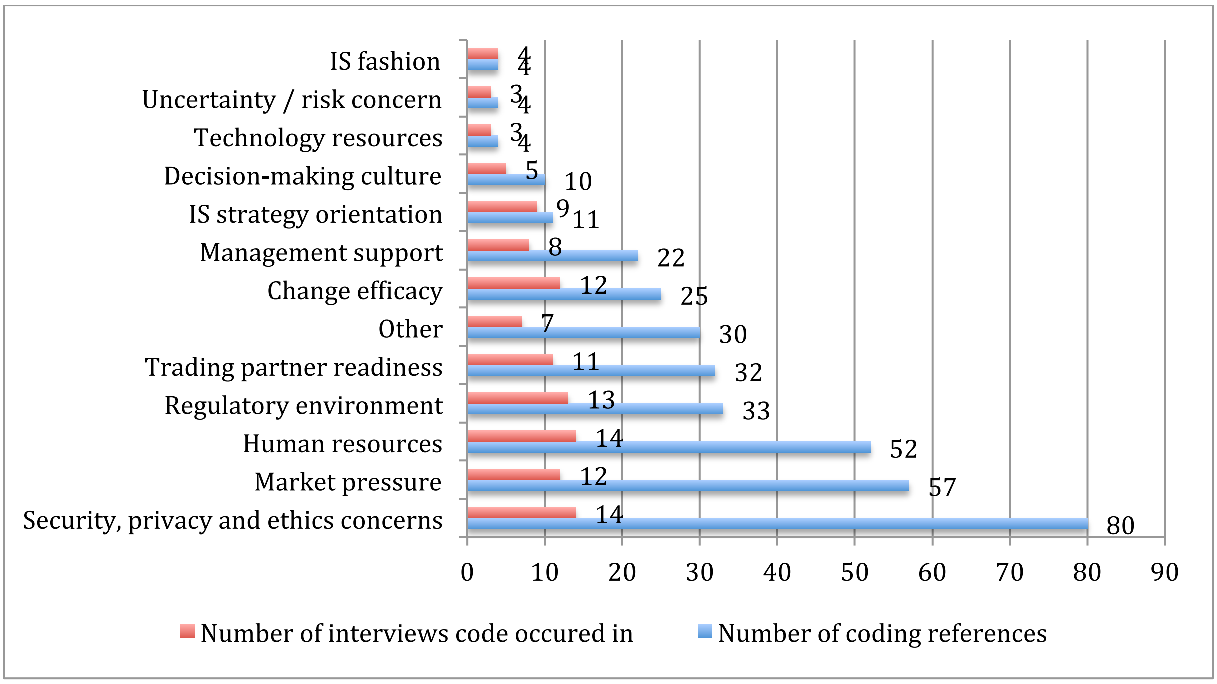

Figure 8 - Overall number of coding references

7.1 Environmental factors influencing big data adoption

As can be seen in Figure 8, the most-mentioned code was **security, privacy and ethics concerns**. This was coded a total of 80 times over all fourteen interviews. However, only one interviewee stated that there were personal concerns regarding ethical aspects of big data projects: "*Big data is an enormous concentration of knowledge and knowledge is power. If you give these tools to the government, a citizen lacks all defense against such a powerful government*" (Int. 4). The same interviewee felt that there was too little attention going to such ethical aspects of big data.

Nine of the interviewees represented management consultancy firms. As a result, while they are the ones that have adopted big data technologies, the data is often of their clients. Therefore, an often-heard response was that the interviewees had no concerns regarding the data. If there were any concerns, it often was about the integrity or security of the data. This relates to missing values in databases or having to correct values with algorithms (Berntzen & Krumova, 2017). However, it also relates to not being sure how data is compiled. For example, firms that use external data sources for analysis said: "*It is easy to make wrong conclusions from correlations. To understand whether a relationship is correlation or causation you have to understand the data and the methods used to compile it*" (Int. 4).

If the concerns were about security and privacy, then almost all interviewees referred to the General Data Protection Regulation (GDPR) enacted in May 2018 in the Netherlands (Netherlands Chamber of Commerce, 2019). The overlap with the **regulatory environment** factor is most likely due to the recent enactment of the law. However, none of the firms see this as a barrier to the use of big data. In the case of using clients' data, often already strict contracts were in place. For all companies, it was a matter of complying. As one interviewee mentioned: "*We were already complying with everything GDPR asked. At the most, it is a delaying factor, because we have to tell our clients how we maintain privacy and security of the data*" (Int. 10).

For one firm, GDPR even accelerated the decision to switch to a new data source: "*GDPR did play a role, I think it accelerated our process to switch to the other data source*" (Int. 3). For firms that maintain a database with consumer data it did have influence, but it did not deter them or discontinued their use of big data. Instead, they had to invest time and money to comply: "*It (GDPR) did have influence. We had to rewrite our terms and conditions and privacy statement (...) But other than that, there are no issues*" (Int. 6). All the firms have adopted big data before GDPR and none have said that it impacted the use to the point of discontinuance. GDPR posed an external incentive that possibly required the changing of routines through dynamic capabilities, but has had little or no influence on the adoption and use of big data.

Another external factor is **market pressure**, which was coded a total of 57 times across 12 interviews. The pressure of competition did not seem to have major influence on the interviewees or them changing their routines. Some did say that they looked at what the competition was doing, but this seemed more in general than specific to big data: "*When expanding our services we do look to the competition. However, I try to reason from our own position and strengths*" (Int. 8). This is a more inward looking stance that corresponds with the RBV rather than the DC approach to react to external signals.

A possible explanation is the **observability** of big data as discussed in the previous chapter. This factor has been coded only twice, but it can explain why firms did not feel pressure from competitors. As one interviewee put it, when asked about big data awareness: "*People do not use big data, because they are unaware of the possibilities*" (Int. 13). If the benefits of an innovation are not demonstrated, the awareness and pressure to follow adoption of an innovation are low (Alshamaila et al., 2013). One of the respondents tried to raise the internal observability of innovations in general by organizing show-and-tell events within the company: "*With a show-and-tell we hope to trigger each other in the possibilities innovations provide*" (Int. 12). This enables the sensing of opportunities, which is an essential first step in learning and developing capabilities.

In comparison to market pressure from competitors, the **trading partner readiness** is more likely to have influence in big data adoption. It does depend on the type of firm whether customers or suppliers influence the decision to use big data. For example, in the consultancy sector customers come to the client with a specific need or question. These clients have selected a firm to align with a possible solution they require: "*People come to us because they have a mountain of data that is left unused*" (Int. 1). However, sometimes clients come to the firm to solve issues relating to a specific technology: "*Unfortunately I am using Hadoop right now for one of my clients. It had its time and half of the time it has to go to maintenance. It really is not a great technology*" (Int. 2). In this sense, clients do somewhat influence the technologies a firm can use.

Some companies partner with other organizations and benefit from these partnerships. One firm uses the data of an external source to complement its consultancy business and generate evidence-based advice. In this sense, they distinguish themselves from their competition: "*You can differentiate yourself from being better than others, thus you are going to look for better methods and data sources (...). We are using the most advanced data source that enables us to work directly from the source material. This leads to better advice and is a new way to beat the competition*" (Int. 3). Another firm has partnered with and uses the platform of Google Cloud. This limits the technologies they can use, but it is explained that it is sufficiently 'open' to allow for external plug-ins. Furthermore: "*We are kept busy more than enough with Google technologies. Focus on Google is us worth a hundred times more than adapting to a different platform for a client's explicit request*" (Int. 10).

It seems that trading partner readiness relates mostly to pressure exerted by clients to develop certain big data competences and capabilities. There are many firms using big data or offering big data solutions. A firm can choose to partner with other organizations or in case of consultancies be dependent on other organizations for data. The method in which firms become aware about big data is the **IS fashion** factor. Information is gathered through clients and suppliers about potentially using big data. The first step is to actually possess data. The second is to either analyze it yourself or go to a third party for support: "*sometimes companies organize hackatons and invite us to do something with their data*". This was only mentioned once, however.

Overall, the environmental factors have little influence on the firms to develop specific capabilities or adopt big data. These findings are summarized in Table 7. The security, privacy and ethics concerns were expected to have no influence on the adoption of big data. However, a distinction between the concerns has to be made, as the concerns regarding security and privacy related to complying with the rules and regulations of The Netherlands. Nevertheless, none of the firms perceived either this factor or the regulatory environment to be an obstacle.

The trading partner readiness was expected to pressure firms in developing competences and adopting big data. However, as a result of the choice available, it is more likely that customers pick these firms based on the existing competences, rather than forcing them to adapt. In some partnerships between firms, the trading partner did dictate the used systems to enhance compatibility.

Due to the transformative properties of big data, it was expected that competitors could pressure firms to adopt big data, to avoid lagging behind. However, the market pressure was not perceived to be of influence in adopting big data. This indicates that the majority of the firms do not see big data as essential for competitiveness yet. As a result, policy makers rather than entrepreneurs perceive the lagging of big data adoption by SMEs in the EU as a problem.

Finally, the IS fashion was expected to have a negative impact on big data adoption if communication channels were lacking. This factor was only mentioned in the context of trading partners, which could influence the decision for certain systems.

Table 7 - Summary of findings on environmental factors

Factor	Expected effects	Findings
Security, privacy and ethics concerns	Concerns on these issues do not influence big data use	Concerns related mostly to regulation and data integrity, but had no effect on big data use
Trading partner readiness	Partners using big data positively influences big data use	Influences the decision to use big data and the choice of tools
Regulatory environment	Regulations negatively influence big data use	GDPR did not hinder big data use
Market pressure	Market pressure has positive influence on big data use	Competitive pressure did not influence the decision to use big data
IS fashion	Lack of communication channels will impact use of big data negatively	Information is gathered through clients and suppliers, which promotes big data use

7.2 Organizational factors influencing big data adoption

The factor **human resources** was coded a total of 52 times in all interviews. This mostly referred to the skills needed for big data. Everyone agreed that working with big data requires a certain skill set. Right now, these are in high-demand, resulting in a premium price for these skills on the labor market (European Commission, 2016a). However, not all firms are clear in the type of skills they need: "*If you look at job descriptions (…) you see that they are looking for the Holy Grail. Someone that is able do everything: a data engineer and data scientist*" (Int. 4). Data engineers are typically involved in programming, while data scientist with the extraction of insights out of data.

For one firm the biggest problem in building an algorithm was finding the right people: "*These people are very expensive and therefore the biggest problem. We had to decide what we wanted to outsource and what we wanted to do ourselves*" (Int. 5). This has to do with the innovation characteristic **complexity**. Behind some of the uses of big data is some difficult math and statistics, which encouraged one participant to follow a specific master's degree (Int. 4).

A person with the rights skills is an important enabler in the use of big data. Larger firms might be better suited to hire these people, but there is also another factor that might hinder SMEs in hiring: "*A data scientist is a researcher in its core. If a firm cannot offer enough challenges, it will have trouble retaining the employee*" (Int. 1). The firm's size does therefore not necessarily dictate whether a firm can use big data. It is likely to be correlated to ability to pay for higher salaries and offer a more challenging job.

The skills gap that is the focus of the overarching project does indeed seem to be hindering successful implementation of big data within SMEs. Small firms simply lack the money to pay the premium for skilled people on the labor market. Nevertheless, the indication that several of the interviewees started experimenting with and learning about big data themselves means that SMEs could very well compensate the skills gap to some extent. It also indicates that dynamic capabilities are in place to add to existing competences and build up organizational knowledge. The choice to strategically support these initiatives are at management level.

This leads to the factor **management support**, which was mentioned 22 times across eight interviews. An important aspect is that this stimulates employees to look around for new technological opportunities. In one firm, the interviewee felt that the 'DNA' of the firm was not tailored towards experimenting with big data. Out of own interest, the interviewee started experimenting with big data and got management interested (Int. 4). More interviewees responded that they saw opportunities for new technologies and started experimenting (Int. 3, 7, 9).

However: "*Like all consultancy firms there is a pressure to be billable. It is likely you have to learn about big data in your own time if there is no payoff in the foreseeable future. After overcoming the initial inertia, I got the interest and support of the management*" (Int. 4). The interest of employees could be a great enabler of fostering innovation and implementing big data. The support of management is therefore instrumental in successful implementation. Management that prioritizes short-term incentives over possible risky investments in employees might hinder successful implementation.

As one interviewee put it: "*Working with an involved CEO of a SME is fantastic, but it could lead to the CEO becoming involved in micro-management and frustrating the process of implementing new technologies*" (Int. 10). It seems that management therefore needs to strike a balance in supporting employees, but do not need to get involved in micro-managing the daily process.

For firms that started in the IT sector, the management support seems to have been successful in leveraging resources for big data implementation. Firms that have started using big data as an addition to their core business did have some issues in gaining full support from the management. For managers it is difficult to identify the opportunities for big data when there is no clear application or goal in mind. However, when employees identify opportunities and want to develop them, management should not stand in their way. While smaller SMEs might not possess the resources to support these employees, the non-hierarchical archetype of SMEs could foster conversation and understanding between employees and management. In turn, this contributes to innovation and successful implementation.

This informal organizational structure was cited as a great advantage in the factor **change efficacy**. It is linked to a culture of flexibility and experimentation, which is something especially start-ups are famous for. Some of these firms were a little older and have added new technologies during their existence, indicating capability for change and sensing opportunities. For others the use of big data was a first time: "*In terms of big data, we are still pioneering*" (Int. 3). Another said: "*We have a very informal culture. It enables us to quickly consult each other and check what is going on. After every project, we evaluate if there are better methods we could have used*" (Int. 6). This openness to new ideas requires a culture that fosters innovation and enables easy transition. Both the management and employees play a crucial role in enabling change through the dynamic capabilities of the firm.

Related to the culture of the firm is the way decisions are made. The factor **decision-making culture** captures whether decisions are made based on evidence. It was only mentioned ten times, but with a focus on the decision-structure of the firm in deciding to adopt big data. For example: "*managers of our clients decide to use big data, but if we ask what for they do not know*" (Int. 2). It indicates that managers might decide to follow the hype surrounding big data, without considering the need or use case. In larger firms, new ideas have to "*go to the board first, which then decides if big data fits the strategy*" (Int. 11). This relates to the path dependency of the RBV, in which prior decisions and assets determine the options that are available to the strategic management of the firm. In this context, it seems that this firm is less dynamic, which results in hindering big data use somewhat by limiting flexibility and management support.

The need to align the strategy of the firm with big data is highlighted by the factor **IS strategy orientation**. Big data has to complement the strategy to be useful. This is not always the case, however: "*some of our clients have a complete mismatch between the data they have and their strategy*" (Int. 11). The collection and analysis of big data have to be done with a clear goal in mind. Some of the interviewed firms used big data specifically to complement their business. For example, utilizing BDA for targeted marketing, creating algorithms or answering geospatial questions. Without this alignment, big data use is not necessarily hindered, but does not represent any value: "*IT has to support the business*" (Int. 8). In turn, dynamic capabilities are needed to better exploit a resource for the firm.

The need for adequate infrastructure is captured in the factor **technology resources**. Different big data technologies exist and the one a firm will use is most likely the result of the history and resources of the firm. This is the case in the previous examples of Hadoop and Google Cloud. However, little was said about having adequate hard- and software for the use of big data. One interviewee did say the following: "*the infrastructure is not a problem, because in theory everyone can use cheap big data technologies in the cloud*" (Int. 2).

Of course, there is the issue of **compatibility**. Three participants mentioned there was an issue of compatibility between the current infrastructure and the new big data infrastructure. They were "*thinking about connecting the different systems*" (Int. 14), but this proved difficult because of the **cost of adoption**: "*the organization of information in firms is costly. Once technologies become cheaper and open standards are common, it is cheaper to exchange information between systems*" (Int. 11). No doubt that cheaper technologies and open standards will enable more firms to start experimenting with big data. However, the cost for personnel was also mentioned as a big hindrance for big data use.

A solution could be to outsource big data problems to third parties. The upside would be twofold: less **uncertainty / risk concern** and an option for **trialability**. As previously mentioned, employees in consultancy firm are pressured to be billable. When an opportunity with big data presents itself, it might pose a great long-term uncertainty of pay-off: "*It* (data science) *is a science and you cannot guarantee the outcome*" (Int. 1). To counter this, big data solutions can be sampled as a proof of concept or minimum viable product (MVP). This decreases both uncertainty and the initial investment requirement. However: "*If a client is unwilling to pay for a proof of concept, they might not be serious. They not only need to feel the pain, but also see the merit of the solution*" (Int. 8).

Furthermore, big data has to have a **relative advantage** compared to other or previous solutions. Some claim it is hype and normal data is sufficient to provide insights (Int. 9). Others do perceive the advantage of big data, because it is not just spreadsheets of data. Interviewee 5 started experimenting with machine learning due to the big database the firm was building: "*We started experimenting to see if we could make it (selecting photos) easier for our customers*" (Int. 5). In turn, big Data could provide a competitive advantage for firms, but four interviewees agreed that small firms without data – "the bakery around the corner" – would never hire a data scientist (Int. 2, 9, 10, 11): "*It is just not feasible. Even if a grocer would need big data, I can only see a situation where it is a one-time only thing*" (Int. 9). In this case, it comes back to the type of data and required analysis whether big data provides a relative advantage or extra effort with no added benefit.

The final factor is **other**, which incorporates statements that did not fall under the other factors. For example, the amount of data a firm possesses is likely an enabler of big data use. One firm has a large database of customer data. It started collecting data around fifteen years ago, but due to new technologies it decided to apply more advanced methods to better utilize its growing database: "*We collected data for years, but we did not have the statistical models we do have now. Now we do more with the same data*" (Int. 6).

Other statements were about the type of firm. Digital SMEs, firms that collect big data themselves and business to consumer (B2C) firms were all mentioned as likely candidates for users of big data: "*We are B2B, so our customer base is small. B2C can have a lot of* *customers that can generate a lot of data*" (Int. 2). This is in line with the business models of large tech companies, where the customers "are the product". In short, if the nature of the firm promotes data collection, it is likely to use big data for the purpose of capability building or as an organizational resource.

Overall, the organizational resources play an important role in the adoption and integration of big data in the firm. The findings on the organizational factors are summarized in Table 8 on the next page and will be briefly discussed here.

The lack of human resources impacts the firm's ability to sense opportunities and fully exploit big data. In this sense, they lack both the skills and ability to acquire them through dynamic capabilities. The acquisition of new skills has to be supported by the management however, as they manage the strategy of the firm and can choose between path-dependent options. These influences were expected and confirmed by the findings.

However, the decision-making culture was found to inhibit the change efficacy and management support factors. Due to the structure of the firm, routines were more difficult to be changed by dynamic capabilities. This is also related to the uncertainty of the payoff of big data use, which prefers short-term certainties to long-term pay-offs.

The technology resources of the firm were expected to inhibit the full exploitation of big data, but were found to be more dependent on the technology resources of the trading partners. Without compatibility between trading partners and the firm, big data transfer was more difficult.

The IS strategy orientation is in theory concerned with exploring options. This is about sensing opportunities that could prove valuable to the firm. In the findings, this factor was interpreted as the alignment between IT and the strategy of the firm, as interviewees mentioned this was crucial for success.

Table 8 - Summary of findings on organizational factors

Factor	Expected effects	Findings
Human resources	Insufficient skills hinders use of big data	Lack of qualified people made it difficult for some firms to fully commit to big data use
Management support	Lack of support from management hinders use of big data	Management plays a crucial role in allocating resources for the adoption and integration of big data.
Technology resources	Insufficient quality of existing infrastructure hinders use of big data	Is related more to the technology of partners and compatibility
Decision-making culture	A risk averse owner hinders use of big data	Hinders big data use by limiting flexibility and management support
Change efficacy	Flexible SMEs are likelier to use big data	A flexible and non-hierarchical structure promote innovation and big data adoption through dynamic capabilities
IS strategy orientation	Exploring options positively influences big data use	Big data use has to be aligned with the strategy of the firm
Uncertainty / risk concern	High uncertainty about big data (platform) reduces the use of big data	Uncertainty about payoff reduces big data use

7.3 Summary of findings

In this chapter, the conceptual framework was used to answer research question four: *what factors influence the adoption of big data by Dutch consultancy SMEs?* Not all firms utilized big data according to the definition of this research, but those that did were successful. The overall findings are summarized in Table 9. The main findings are that a lack of qualified people with the necessary expertise can really hinder the use of big data. Due to the complexity and high costs, the uncertainty about the payoff and need for big data increases. To counter this, trialability of big data could be used to minimize risk.

Furthermore, management has to support employees that experiment with big data by allocating time and resources. They should also consider how big data fits in the overall strategy of the firm. The strength of SMEs is their flexibility and decision-making culture, which enables them to adapt quickly. However, with bureaucratic procedures big data projects might never take root in the organization. Finally, the lack of standardization hinders the ability to exchange data between sources.

Table 9 - Summary of findings

Nr	Factor	Findings
1	Human resources	Lack of qualified people made it difficult for some firms to fully commit to big data use
2	Management support	Management plays a crucial role in allocating resources for the use of big data.
3	Technology resources	Is related more to the technology of partners and compatibility
4	Decision-making culture	Hinders big data use by limiting flexibility and management support
5	Change efficacy	A flexible and non-hierarchical structure promote innovation and big data use
6	IS strategy orientation	Big data use has to be aligned with the strategy of the firm
7	Security, privacy and ethics concerns	Concerns related mostly to regulation and data integrity, but had no effect on big data use
8	Trading partner readiness	Influences the decision to use big data and the choice of tools
9	Regulatory environment	GDPR did not hinder big data use
10	Uncertainty / risk concern	Uncertainty about payoff reduces big data use
11	Market pressure	Competitive pressure did not influence the decision to use big data
12	IS fashion	Information is gathered through clients and suppliers, which promotes big data use
13	Relative advantage	It depends on the case whether big data offers an advantage compared to other solutions
14	Compatibility	Lack of standardization hinders the connection of multiple data sources
15	Complexity	Big data is complex, which influences the required skills needed
16	Trialability	MVP or proof of concept lowers the uncertainty and risk involved
17	Observability	Lack of observability influences perceived competitive pressure
18	Cost of adoption	Cost of personnel and technology hinder big data use
19	Other	Amount of data positively affects big data use

As described in chapter 4, the conceptual framework contained factors that were found to be of importance in the literature on big data use. The factors themselves are not necessarily specific to big data and could refer to IT use in general as well. Based on the empirical investigation, the factors were not shown to contain a different meaning in the context of big data. These results do bring back the question whether big data is truly a driver for competitive advantage and as unique as claimed, as issues similar to traditional IT investment occur in this setting. Based on this research, big data might not be so unique after all.

The factors that were found to affect big data adoption are shown schematically in Figure 10. It provides an overview of the found relationships between the factors, indicating that the proposed research model is too simplistic for this complex situation. For example, change efficacy fulfills the role of dynamic capabilities and many of the IC are found to influence in a mediating relationship. Nevertheless, it visualizes the most important findings of Table 9.

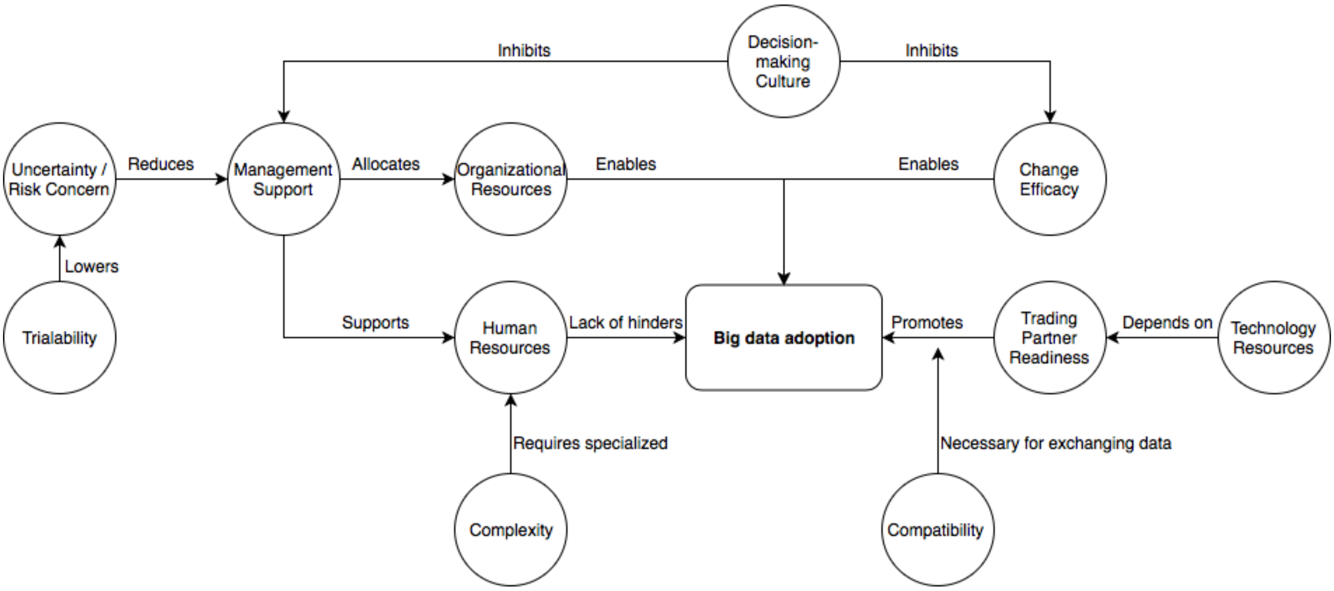

Figure 9 - Schematic overview of the relationships

Chapter 8 – Discussion and Conclusion

Four research questions were formulated to get a better understanding. The first research question was answered in chapter 2 and elaborated on the definition of big data. The aim of the second research question was to find why big data is a unique resource and frame it in the context of creating a competitive advantage. This added to the understanding of how big data is different from traditional IT. The third research question was about determining the challenges SMEs face in the adoption of big data. This question was answered with the conceptual framework in chapter 4. The fourth research question aimed to understand how these factors influence the adoption of big data in practice. This was answered with the conceptual model and qualitative interviews, as described in chapters 4 and 5. The outcomes of these research questions are summarized in this chapter to provide an answer to the main question:

What makes big data unique and how do organizational resources impact its adoption by Dutch small and medium consultancy SMEs?

The contribution of this thesis is on the topic of big data use by SMEs. The aim is to understand what big data is and why SMEs are lagging behind in using it. A lot has been written on the subject, but knowledge of reality in the Dutch context is lacking. First, a conclusion per chapter is given. Then, the main research question is answered. This is followed by a discussion on the research, its limitations and suggestions for future research.

8.1 The definition of big data

The first research question: *what is big data?* was the focus in the second chapter. I found that a consensus on a definition for big data is lacking and that several approaches to defining the concept exist. The first approach defines big data as being high in volume, variety and velocity. This means that big data is large in size, consists of different types of data like audio and text and is gathered or generated very quickly. This definition is broad and open for interpretation.

The other approach is defining big data by the type of technologies that are needed to process it. This defines big data by requiring new forms of processing, as traditional database management tools are unable to cope. The downside of this definition is that what constitutes traditional is time-dependent. The whole concept of big data is therefore difficult to pin down. To guide this thesis, big data is defined as being: *high volume, high velocity or high variety information assets*.

8.2 Big data as a unique resource

The second research question was: *why is big data a unique resource and how does it differ from traditional IT in Dutch consultancy SMEs?* This question was part theoretical and part empirical in nature. The goal of this question is to frame big data in the context of creating a competitive advantage and conceptualizing it as a VRIN resource in the RBV. The first part of the research question is answered by examining the characteristics of big data through the lens of the RBV. In it, firms are conceptualized as bundles of resources. These resources have certain characteristics – value, rareness, imitability and non-substitutability – that determine how likely they are to lead to a competitive advantage.

The characteristics of big data make it difficult to imitate or substitute. In turn, value is generated from insights that can lead to process optimization, innovation and IT capabilities for the exploitation of big data. It was expected that this could be where the real uniqueness of big data lies, as it facilitates the development of capabilities and knowledge, which enables firms to better leverage resources for the creation of competitive advantages.

The empirical part investigated the reality of big data application and whether it is truly different from traditional IT. The firms ranged wildly in their definition of big data. In total, nine firms did utilize big data according to the definition of this research. However, none combined all aspects of the 3V model. Volume was the most often cited, but some firms also included the technologies they used in the definition of big data. As discussed in chapter 2, tools do not define what big data is, but seeing the development of specialized technologies for big data does warrant an extension of the definition to include the complexity of big data. This would define big data as *either high volume, velocity or variety information assets that are difficult to handle*.

Based on this research, however, there is inconclusive evidence that big data is truly different from traditional IT. In the sample, big data was mostly seen as a product or resource in itself, not a facilitator of higher order capabilities. Furthermore, as is discussed in the next section, the factors found to influence the adoption of big data are similar to other IT adoption processes. If big data is truly different, it could have required a completely different approach.

8.3 Challenges of big data adoption

The third research question is *what are the potential challenges of big data adoption?* It was answered in chapter 4 by performing a literature review and constructing a conceptual framework. It is based on the article of Sun et al. (2018). This article performed a systematic literature review to find all the factors of influence in the organizational adoption of big data.

Factors were divided in characteristics relating to the innovation, organization and environment. The innovation characteristics were excluded from the framework, as the main focus is on the organizational and environmental relations. Sun et al. (2018) described 26 factors, some of which were left out resulting in a conceptual framework with 19 factors (see Table 10).

According to the literature, these factors play a role in the decision to adopt and utilize big data. Whether a factor is of influence is dependent on the context of the particular SME and technology that is being studied. As a result, this list provides an overview and a theoretical generalization of what was found of importance in prior studies. A factor that could have been a major barrier in once case, could have been of no consequence at all in another. However, it did enable the formulation of expected effects. This leads to the next research question.

Table 10 - Potential challenges to the adoption of big data

Nr	Cate gory	Factor	Expected effects
1	OC	Human resources	Insufficient skills hinders use of big data
2	OC	Management support	Lack of support from management hinders use of big data
3	OC	Technology resources	Insufficient quality of existing infrastructure hinders use of big data
4	OC	Decision-making culture	A risk averse owner hinders use of big data
5	OC	Change efficacy	Flexible SMEs are likelier to use big data
6	OC	IS strategy orientation	Exploring options positively influences big data use
7	EC	Security, privacy and ethics concerns	Concerns on these issues do not influence big data use
8	EC	Trading partner readiness	Partners using big data positively influences big data use
9	EC	Regulatory environment	Regulations negatively influence big data use
10	EC	Uncertainty / risk concern	High uncertainty about big data (platform) reduces the use of big data
11	EC	Market pressure	Market pressure has positive influence on big data use
12	EC	IS fashion	Lack of communication channels will impact use of big data negatively

8.4 Factors influencing big data use in Dutch SMEs

The fourth research question is: *what factors influence the adoption of big data by Dutch consultancy SMEs?* It was answered with the use of interviews in chapter 7. This method was chosen due to the newness of big data and lack of data on the subject. The interviews provided insights in individual cases and how a firm adopted the use of big data. Interviews were the best method to reconstruct the context in which big data was adopted and used.

The trends that emerged provided a starting point for understanding which factors were of importance in practice in enabling or hindering the adoption and use of big data. Some factors could be enablers for one firm, but perceived as a major barrier for firms that lacked this factor. For example, change efficacy could have helped the use of big data in a small firm with a non-hierarchical structure and culture. In firms that lacked this structure and openness towards change, it felt that a lot of inertia had to be overcome in order to adapt to big data within the firm.

The biggest factor hindering big data use was human resources. Interviewees cited that acquiring people with the right set of skills is the most difficult thing to do for an SME. The skills that are required to work with big data are not unique, but are related to data scientists and engineers in general. The problem is that people with these qualifications are in very high demand. In turn, SMEs lack the money and sometimes challenging and diverse projects to attract and retain these people. Money was also cited as a barrier in acquiring technologies, while some interviewees contested this by stating that cloud platforms are cheap and open enough to do almost all business processes in.

Surprisingly, many factors were found to be of little or no perceived influence. The ethical side of the use of big data was mentioned only once. Most of the concerns went to the integrity, privacy and security of the data. However, this did not deter firms from adopting big data. The security and privacy of the data was also an issue in partnerships between organizations. However, GDPR seemed to be an enabler in fostering trust rather than being a barrier.

Cooperation and partnerships are common, especially since most of the interviewed firms belong to the consultancy branch. Competition between these firms did not affect the adoption of certain technologies. Interviewees experienced the pressure from clients mostly in the form of finding them, acquiring projects and securing data. In B2B firms the clients would often come to the firm because of a certain need, not with the request to use a specific technology for the solution.

Another important finding is that no factors emerged that were specific to big data. All other factors, though emerging from a literature study on big data, could apply to the broader context of IT. The empirical research in this thesis did not change the meaning of these factors in the particular context of big data that would not also apply to more general use of

IT. This call into question whether big data is truly unique, compared to the prior history of IT investments.

The findings of the interviews are summarized in Figure 11. This provides an overview of the organizational resources and dynamic capability, or change efficacy, and their relationships. It is different from the proposed research model, because reality is messier than theory. The findings did not support a strict separation in organizational resources and dynamic capabilities. Furthermore, the innovation characteristics were found to moderate in some way certain organizational resources.

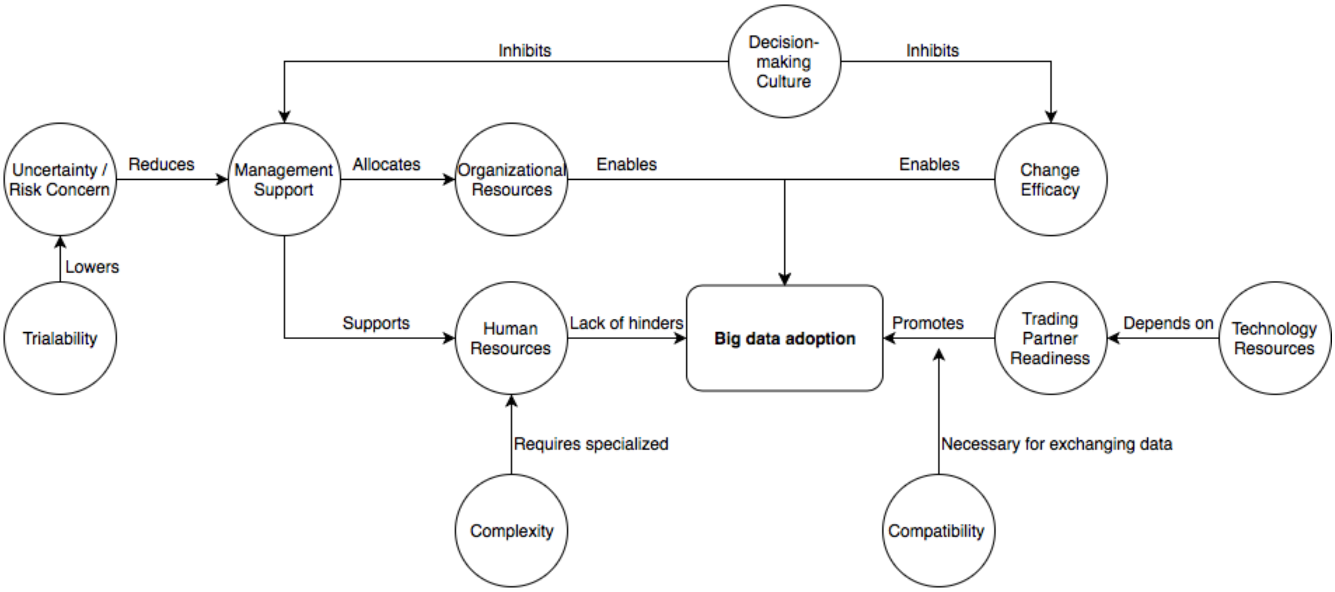

Figure 10 - Overview of the relationships

8.5 Implications

The main research question is *what makes big data unique and how do organizational resources impact its adoption by Dutch small and medium consultancy SMEs?* It will be answered based on the previous findings. To start, this research did conceptualize big data as a unique resource based on its characteristics. In theory, it could contribute to the creation of a competitive advantage in two ways.

First, by being a unique organizational resource. Second, through the fostering of learning and development of dynamic capabilities, which enables a better exploitation of the assets of the firm. However, in practice it is used mostly as a regular organizational resource, not too much different in use and application than regular IT. This brings nuance to the claim that big data is transformative. At least for now in the Dutch consultancy SMEs, it is not.

Second, the overview of all the factors of importance have to be kept in mind to answer which organizational resource impact big data adoption. Environmental factors were found to be of much less importance than organizational resources. In contrast, dynamic capabilities did not emerge from the interviews, other than the factor change efficacy. This reinforces the idea that the adoption of big data is currently an internal affair, without much pressure from the environment. As a result, the need for big data adoption is perceived much more pressing for policy makers than entrepreneurs.

The biggest barriers to the adoption and use of big data by SMEs are skills, money and the availability of data. Due to the lack of skilled – and for SMEs affordable – people in the market, some firms started training people in-house. This is a viable solution for some firms, but does require money and the culture to enable it. Therefore, on the policy level, more people need to be trained to become data scientists and engineers. This is how policy makers could perceive the lag of big data adoption as a problem. Due to the high premium the market provides for high-skilled people, students are more likely to choose educational programs that train them in programming, statistics, coding and math.

However, steps could be taken to increase the support for and diversity of these students, to include people with different (academic) backgrounds that might think differently on issues like ethics. Engineers are likely to perceive ethical concerns differently than social scientists. This also highlights that closing the skills gap is about more than educating more data scientists for the future workforce. Students in all disciplines have to acquire digital skills at some stage in their training and reskilling current employees could provide a short-term solution for the shortage.

On the firm level, managers need to be aware of the reason for adopting big data. The results of big data analytics can be unknown and may very well not result in immediate payoff. However, this does not mean that there should not be time and support to experiment with these technologies. It might be easier to formulate expectations if a firm possesses own data sources that could be used to optimize processes or increase efficiency. Otherwise, partnering with other organizations to create more data could prove valuable.

However, one should remain vigilant on what data can and cannot do. Basing decisions on correlations without understanding the underlying relationships is unlikely to be a sustainable way of doing business. A solid understanding of what data represents and what conclusions can or cannot be based on analysis remains important.

8.6 Limitations

Academic literature has mostly focused on studying improving use and adoption of innovations rather than preventing diffusion of unwanted innovations (Rogers, 2003). This pro-innovation bias is also implicit in this research, as it is assumed that big data will play an increasingly important role. However, as big data has still to be diffused successfully to all SMEs, this research is still valuable because it does not focus on an already successfully diffused innovation.

A second bias is the recall problem. This is the tendency of looking back in qualitative research and thereby not accurately recalling events that have happened. The less time between the research and the object of study, the less likely a recall bias is introduced. Studying a recent innovation like big data and giving interviewers better training can compensate the recall bias (Bryman, 2012). Furthermore, some of the interviewees had just started using big data and were still finding new uses for them. This meant that there was little time between the adoption decision and when the interview took place.

One of the limitations of this research is that it looks at the use of big data in companies that have successfully adopted the innovation. While this could introduce the recall problem, it does offer the opportunity to study successful companies. In addition, due to the currently limited diffusion of big data within SMEs these firms are the only option for in-depth analysis. Firms currently in the process of adopting big data are often hesitant to share knowledge about the process, in fear that their competition might get an edge.

A related limitation concerns the people that I interviewed. They are part of the firm, which gives them a unique perspective that could also introduce a bias. Another related limitation is in the structure of the interviews. Some factors were explicitly questioned, which influences the amount of coded references. As a result, factors that emerged during the interviews are coded less frequently. However, if the interviewees did not bring up these factors, it is assumed that they are not of interest in their particular context.

Another limitation of this thesis is the focus on a specific geographical area. While this is necessary to provide sufficient depth and relevance, it does ignore the reality that factors are dependent on cultural and geographical context. What might be an enabler in The Netherlands could play a very different role in another country or region. Furthermore, The Netherlands is known for its focus on rules and regulations. In this context, GDPR might indeed be of little influence. However, in countries with less privacy regulations already in place it might be a big hurdle to take. Similarly, The Netherlands might be an attractive country for skilled expats, lessening the problem of a lack of skilled people while exacerbating the brain drain in other places.

8.7 Suggestions for future research

The findings and limitations of this research provide opportunities for future research. The nature of this thesis is exploratory. As a result, many firms were approached, but mostly consultancy firms and firms active in IT services were interested in participating. Future research might study barriers to big data use in more detail by focusing on specific sectors or types of firms. For example, one of the interviewees mentioned that B2C firms might benefit more from using big data than B2B firms. Future research could study this in The Netherlands, but other countries could provide valuable insights as well. This would enable a cultural comparison between countries.

The hype currently surrounding big data and other digital technologies could indicate that there is a technology push going on. This hype could persuade management of some firms to adopt these technologies, without having an actual need. Future research could study how important the use of big data is in SMEs. An often-heard argument is that the impact of big data is bigger in larger firms. Smaller firms could benefit more from investing in other things than big data. Especially considering the findings that the use of big data does not differ significantly from other IT projects. Understanding what the needs are for SMEs and how big data could provide a solution would be valuable insights.

Finally, the proposed research model focused on the use of organizational resources and the development of dynamic capabilities. It is assumed that these lead to an increase in firm performance or a competitive advantage. The exact pathway of creating such a competitive advantage based on big data requires further study on the link between big data adoption and competitive advantage.

References

Agrawal, K. P. (2015). Investigating the determinants of Big Data Analytics (BDA) adoption in emerging economies. *Academy of Management Proceedings*, *2015*(1), 11290–11290. https://doi.org/10.5465/AMBPP.2015.11290abstract

Al-Qirim, N., Tarhini, A., & Rouibah, K. (2017). Determinants of Big Data Adoption and Success. In *Proceedings of the International Conference on Algorithms, Computing and Systems - ICACS '17* (pp. 88–92). https://doi.org/10.1145/3127942.3127961

Alshamaila, Y., Papagiannidis, S., & Li, F. (2013). Cloud computing adoption by SMEs in the north east of England: A multi-perspective framework. *Journal of Enterprise Information Management*, *26*(3), 250–275.

Antlová, K. (2009). Motivation and barriers of ICT adoption in small and medium-sized enterprises. *EM Ekonomie a Management*, *12*(2), 140–155. Retrieved from http://proquest.umi.com/pqdlink?did=1743680191&sid=2&Fmt=7&clientId=45608&RQT=309&VName=PQD

Antonioli, D., & Della Torre, E. (2016). Innovation adoption and training activities in SMEs. *International Journal of Human Resource Management*, *27*(3), 311–337. https://doi.org/10.1080/09585192.2015.1042901

Anwar, S., & Khaki, S. (2017). Big Data : Concepts, Challenges and Solutions. Retrieved December 15, 2018, from http://bigdata-tech.blogspot.com/p/big-data-definitions-and-concepts.html

Ardagna, C. A., Ceravolo, P., & Damiani, E. (2016). Big data analytics as-a-service: Issues and challenges. *2016 IEEE International Conference on Big Data (Big Data)*, 3638–3644. https://doi.org/10.1109/BigData.2016.7841029

Awa, H. O., Ukoha, O., & Emecheta, B. C. (2012). Integrating TAM and TOE Frameworks and Expanding their Characteristic Constructs for E-Commerce Adoption by SMEs. In *Information Science & IT Education Conference (InSITE)* (pp. 570–588).

Barney, J. B., Ketchen, D. J., & Wright, M. (2011). The future of resource-based theory: Revitalization or decline? *Journal of Management*, *37*(5), 1299–1315. https://doi.org/10.1177/0149206310391805

Benhabib, J., & Spiegel, M. M. (2005). Human Capital and Technology Diffusion. In *Handbook of Economic Growth* (Vol. 1, pp. 935–966). https://doi.org/10.1016/S1574-0684(05)01013-0

Berntzen, L., & Krumova, M. (2017). Big Data from a Business Perspective. In M. Themistocleous & V. Morabito (Eds.), *Information Systems* (pp. 119–127). Coimbra, Portugal: Springer. Retrieved from https://link.springer.com/content/pdf/10.1007%2F978-3-319-65930-5.pdf

Boothby, D., Dufour, A., & Tang, J. (2010). Technology adoption, training and productivity performance. *Research Policy*, *39*(5), 650–661. https://doi.org/10.1016/j.respol.2010.02.011

Bryman, A. (2012). *Social Research Methods* (4th ed.). Oxford University Press.

Brynjolfsson, E. (1993). The productivity paradox of information technology. *Communication of the ACM*, *3*(12), 67–77. https://doi.org/10.1111/j.1468-

2370.2008.00250.x
Brynjolfsson, E., & Hitt, L. M. (1998). Beyond the productivity paradox. *Communications of the ACM*, *41*(8), 49–55. https://doi.org/10.1145/280324.280332
Brynjolfsson, E., & McElheran, K. (2016). The Rapid Adoption of Data-Driven Decision-Making. *American Economic Review*, *106*(5), 133–139.
Caesarius, L. M., & Hohenthal, J. (2018). Searching for big data: How incumbents explore a possible adoption of big data technologies. *Scandinavian Journal of Management*, *34*(2), 129–140. https://doi.org/10.1016/j.scaman.2017.12.002
Cavanillas, J. M., Curry, E., & Wahlster, W. (2016). *New horizons for a data-driven economy: A Roadmap for Usage and Exploitation of Big Data in Europe*. (J. M. Cavanillas, E. Curry, & W. Wahlster, Eds.), *New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe*. Springer Open. https://doi.org/10.1007/978-3-319-21569-3_16
Cedefop. (2018). *Insights into skill shortages and skill mismatch: learning from Cedefop's European skills and jobs survey*. Luxembourg: Publications Office. Retrieved from http://data.europa.eu/doi/10.2801/645011%0AA
Centraal Bureau voor de Statistiek. (2017). ICT-gebruik bij bedrijven; bedrijfsgrootte, 2017. Retrieved June 22, 2018, from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83869NED/table?ts=1528729057223
Centraal Bureau voor de Statistiek. (2018). Bedrijven, bedrijfstak. Retrieved December 4, 2018, from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/81589ned/table?ts=1543924287951
Cesário, M., Fernandes, S., Jesus, B., & Barata, J. M. (2015). Sources of innovation: The case of portuguese consultancy sector. *Journal of Technology Management and Innovation*, *10*(3), 44–52. https://doi.org/10.4067/S0718-27242015000300005
Chae, H., Koh, C. E., & Prybutok, V. R. (2014). Information Technology Capability and Firm Performance: Contradictory Findings and their possible causes. *MIS Quarterly*, *38*(1), 305–326.
Christensen, C. M., Wang, D., & van Bever, D. (2013). Consulting on the cusp of disruption. *Harvard Business Review*, (OCT). https://doi.org/10.1002/sej.121
Coleman, S., Göb, R., Manco, G., Pievatolo, A., Tort-Martorell, X., & Reis, M. S. (2016). How Can SMEs Benefit from Big Data? Challenges and a Path Forward. *Quality and Reliability Engineering International*, *32*(6), 2151–2164. https://doi.org/10.1002/qre.2008
Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.). SAGE Publications. https://doi.org/10.1007/s13398-014-0173-7.2
Cukier, K., & Mayer-Schoenberger, V. (2013). Rise of Big Data: How it's Changing the Way We Think about the World. *Foreign Affairs*, *92*(3), 28–40. https://doi.org/10.1007/s13398-014-0173-7.2
Davies, R. (2016). Big Data and Data Analytics: The potential for innovation and growth. *European Parliamentary Research Service*, (September). https://doi.org/10.4018/IJSWIS.2017010101
DG GROW. (2018). What is an SME? Retrieved April 5, 2018, from http://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition_nl

Dosi, G., Moschella, D., Pugliese, E., & Tamagni, F. (2015). Productivity, market selection, and corporate growth: comparative evidence across US and Europe. *Small Business Economics*, *45*(3), 643–672. https://doi.org/10.1007/s11187-015-9655-z

Easterby-Smith, M., & Prieto, I. M. (2008). Dynamic capabilities and knowledge management: An integrative role for learning? *British Journal of Management*, *19*(3), 235–249. https://doi.org/10.1111/j.1467-8551.2007.00543.x

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic Capabilities: What Are They? *Strategic Management Journal*, *21*(10/11), 1105–1121. https://doi.org/10.1002/smj.176

European Commission. (2016a). *Analytical underpinning for a New Skills Agenda for Europe*. Retrieved from http://ec.europa.eu/social/BlobServlet?docId=15691&langId=en

European Commission. (2016b). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A new skills agenda for Europe*. *Com/2016/0381*. Brussels. Retrieved from http://www.ipex.eu/IPEXL-WEB/dossier/document/COM20160381.do

Executive Agency for SMEs. (2017). Supporting specialised skills development: big data, Internet of things and cybersecurity for SMEs. Retrieved March 21, 2018, from https://ec.europa.eu/easme/en/tender/10016/supporting-specialised-skills-development-big-data-internet-things-and-cybersecurity

Ferneley, E., & Bell, F. (2006). Using bricolage to integrate business and information technology innovation in SMEs. *Technovation*, *26*(2), 232–241. https://doi.org/10.1016/j.technovation.2005.03.005

Fink, D. (1998). Guidelines for the Successful Adoption of Information Technology in Small and Medium Enterprises. *International Journal of Information Management*, *18*(4), 243–253. https://doi.org/10.1016/S0268-4012(98)00013-9

Freel, M. S. (2005). Patterns of innovation and skills in small firms. *Technovation*, *25*(2), 123–134. https://doi.org/10.1016/S0166-4972(03)00082-8

Freese, J., & Peterson, D. (2017). Replication in Social Science. *Annual Review of Sociology*, (43), 147–165. https://doi.org/10.1146/annurev-soc-060116-053450

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Google Trends. (2018). Big Data search interest from 2004 to now. Retrieved May 28, 2018, from https://trends.google.com/trends/explore?date=all&q=big data

Gray, C. (2006). Absorptive capacity, knowledge management and innovation in entrepreneurial small firms. *International Journal of Entrepreneurial Behavior & Research*, *12*(6), 345–360. https://doi.org/10.1108/09574090210806333

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, *70*, 308–317. https://doi.org/10.1016/j.jbusres.2016.08.004

Gupta, P., Seetharaman, A., & Raj, J. R. (2013). The usage and adoption of cloud computing by small and medium businesses. *International Journal of Information Management*, *33*(5), 861–874. https://doi.org/10.1016/j.ijinfomgt.2013.07.001

Hammer, C. L., Kostroch, D. C., & Quirós, G. (2017). Big Data: Potential, Challenges, and

Statistical Implications. *IMF Staff Discussion Note*, *17*(6), 41. https://doi.org/9781484310908

Hartmann, P. M., 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*, *36*(10), 1382–1406. https://doi.org/http://dx.doi.org/10.1108/MRR-09-2015-0216

Hausman, A. (2005). Innovativeness among small businesses: Theory and propositions for future research. *Industrial Marketing Management*, *34*(8), 773–782. https://doi.org/10.1016/j.indmarman.2004.12.009

Hoppe, H. C. (2002). The timing of new technology adoption: Theoretical models and empirical evidence. *Manchester School*, *70*(1), 56–76. https://doi.org/10.1111/1467-9957.00283

IDC, & The Lisbon Council. (2018a). *First Report on Facts and Figures: Updating the European Data Market Study Monitoring Tool*.

IDC, & The Lisbon Council. (2018b). *How the power of data will drive EU economy: The European Data Market Monitoring Tool Report*. Retrieved from http://datalandscape.eu/sites/default/files/report/EDM_D2.2_First_Report_on_Policy_Conclusions_20.04.2018.pdf

Khalifa, M., & Davison, R. M. (2006). SME adoption of IT: The case of electronic trading systems. *IEEE Transactions on Engineering Management*, *53*(2), 275–284. https://doi.org/10.1109/TEM.2006.872251

Kraaijenbrink, J., Spender, J. C., & Groen, A. J. (2011). The Resource-Based View: A Review and Assessment of Its Critiques. *Silva Balcanica*, *12*(1), 37–46. https://doi.org/10.1177/0149206309350775

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, *34*(3), 387–394. https://doi.org/10.1016/j.ijinfomgt.2014.02.002

Laney, D. (2001). META Delta. *Application Delivery Strategies*, *949*(February 2001), 4. https://doi.org/10.1016/j.infsof.2008.09.005

Lee, J., Kao, H. A., & Yang, S. (2014). Service innovation and smart analytics for Industry 4.0 and big data environment. *Procedia CIRP*, *16*, 3–8. https://doi.org/10.1016/j.procir.2014.02.001

Lefebvre, L. A., Lefebvre, É., & Harvey, J. (1996). Intangible assets as determinants of advanced manufacturing technology adoption in SME's: toward an evolutionary model. *IEEE Transactions on Engineering Management*, *43*(3), 307–322. https://doi.org/10.1109/17.511841

Levy, M., Powell, P., & Worrall, L. (2005). Strategic Intent and E-Business in SMEs. *Information Resources Management Journal*, *18*(4), 1–20. https://doi.org/10.4018/irmj.2005100101

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *Journal of Strategic Information Systems*, *24*(3), 149–157. https://doi.org/10.1016/j.jsis.2015.08.002

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*.

McKinsey Global Institute. https://doi.org/10.1080/01443610903114527
Mayring, P. (2000). Qualitative Content Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(2), 1. https://doi.org/10.1111/j.1365-2648.2007.04569.x
Mazzei, M. J., & Noble, D. (2017). Big data dreams: A framework for corporate strategy. *Business Horizons*, *60*(3), 405–414. https://doi.org/10.1016/j.bushor.2017.01.010
McAfee, A., & Brynjolfsson, E. (2012). Big data: The Management Revolution. *Harvard Business Review*, *90*(10), 61–67. https://doi.org/00475394
Miastkowski, S. (1998). When it comes to hard drives, size is everything. Retrieved August 23, 2018, from http://edition.cnn.com/TECH/computing/9806/12/harddrive.idg/
Michael, K., & Miller, K. W. (2013). Big data: New opportunities and new challenges. *Computer*, *46*(6), 22–24. https://doi.org/10.1109/MC.2013.196
Miller, H. G., & Mork, P. (2013). From data to decisions: A value chain for big data. *IEEE Computer Society*, *15*(1), 57–59. https://doi.org/10.1109/MITP.2013.11
Miller, S. (2014). Collaborative Approaches Needed to Close the Big Data Skills Gap. *Journal of Organization Design*, *3*(1), 26. https://doi.org/10.7146/jod.9823
NESSI. (2012). Big Data: A New World of Opportunities. *NESSI White Paper*, (December), 25.
Netherlands Chamber of Commerce. (2019). The General Data Protection Regulation (GDPR) step-by-step. Retrieved February 19, 2019, from https://business.gov.nl/running-your-business/business-management/administration/the-gdpr-step-by-step/
Newbert, S. L. (2008). Value, Rareness, Competitive Advantage and Performance: A conceptual-level empirical investigation of the resource-based view of the firm. *Strategic Management Journal*, *29*, 745–768. https://doi.org/10.1002/smj
Obeidat, B. Y., Al Suradi, M. M., Masa'deh, R., & Tarhini, A. (2016). The impact of knowledge management on innovation: An Empirical study on Jordanian consultancy firms. *Management Research Review*, *39*(10), 1214–1238. https://doi.org/10.1108/09574090910954864
Oliner, S. D., & Sichel, D. E. (2000). The Resurgence of Growth in the Late 1990s: Is Information Technology the Story? *Journal of Economic Perspectives*, *14*(4), 3–22. https://doi.org/10.2139/ssrn.233139
Oliveira, T., Thomas, M., & Espadanal, M. (2014). Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. *Information and Management*, *51*(5), 497–510. https://doi.org/10.1016/j.im.2014.03.006
Porter, M. E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. The Free Press. https://doi.org/10.1182/blood-2005-11-4354
Porter, M. E., & Heppelmann, J. E. (2014). How Smart, Connected Products Are Transforming Competition. *Harvard Business Review*, (November), 1–48.
Porter, M. E., & Heppelmann, J. E. (2015). How Smart, Connected Products Are Transforming Companies. *Harvard Business Review*, (October), 1–38.
Press, G. (2013). A Very Short History Of Big Data. Retrieved August 23, 2018, from https://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/#4c148d565a18

Priem, R. L., & Butler, J. E. (2001). Is the Resource-Based "View" a Useful Perspective for Strategic Management Research? *The Academy of Management Review*, *26*(1), 22–40.

Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, *38*(1), 187–195. https://doi.org/10.1016/j.ijinfomgt.2017.07.008

Ramanathan, R., Philpott, E., Duan, Y., & Cao, G. (2017). Adoption of business analytics and impact on performance: a qualitative study in retail. *Production Planning and Control*, *28*(11–12), 985–998. https://doi.org/10.1080/09537287.2017.1336800

Rashid, M. A., & Al-Qirim, N. A. (2001). E-Commerce Technology Adoption Framework by New Zealand Small to Medium Size Enterprises. *Research Letters of the Information, Mathematics and Science*, *2*, 63–70.

Raymond, L., & Uwizeyemungu, S. (2007). A profile of ERP adoption in manufacturing SMEs. *Journal of Enterprise Information Management*, *20*(4), 487–502. https://doi.org/http://dx.doi.org/10.1108/MRR-09-2015-0216

Real, J. C., Leal, A., & Roldán, J. L. (2006). Information technology as a determinant of organizational learning and technological distinctive competencies. *Industrial Marketing Management*, *35*(4), 505–521. https://doi.org/10.1016/j.indmarman.2005.05.004

Riggins, F. J., & Wamba, S. F. (2015). Research directions on the adoption, usage, and impact of the internet of things through the use of big data analytics. In *Proceedings of the Annual Hawaii International Conference on System Sciences (HICSS)*. https://doi.org/10.1109/HICSS.2015.186

Roden, S., Nucciarelli, A., Li, F., & Graham, G. (2017). Big data and the transformation of operations models: A framework and a new research agenda. *Production Planning and Control*, *28*(11–12), 929–944. https://doi.org/10.1080/09537287.2017.1336792

Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). New York: Free Press.

Russom, P. (2011). Big Data Analytics. *TDWI Best Practices Report - White Paper*, (Fourth Quarter), 38.

Schwab, K., Sala-i-Martín, X., & World Economic Forum. (2017). *The Global Competitiveness Report*. (K. Schwab, Ed.). Geneva, Switzerland. https://doi.org/92-95044-35-5

Sen, D., Ozturk, M., & Vayvay, O. (2016). An Overview of Big Data for Growth in SMEs. *Procedia - Social and Behavioral Sciences*, *235*(October), 159–167. https://doi.org/10.1016/j.sbspro.2016.11.011

Shah, S., Soriano, C. B., & Coutroubis, A. D. (2017). Is Big Data for Everyone ? The Challenges of Big Data Adoption in SMEs. In *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)* (pp. 803–807). https://doi.org/10.1109/IEEM.2017.8290002

Shan, S., Luo, Y., Zhou, Y., & Wei, Y. (2018). Big data analysis adaptation and enterprises' competitive advantages : the perspective of dynamic capability and resource-based theories. *Technology Analysis & Strategic Management Strategic Management*, 1–15. https://doi.org/10.1080/09537325.2018.1516866

Sila, I. (2013). *Factors affecting the adoption of B2B e-commerce technologies*. *Electronic Commerce Research* (Vol. 13). https://doi.org/10.1007/s10660-013-9110-7

Simon, P. W. and. (2013). *Too Big to Ignore : The Business Case for Big Data*.

https://doi.org/10.1017/CBO9781107415324.004
Stonebraker, M. (2012). What Does “big Data” Mean? Retrieved August 23, 2018, from https://cacm.acm.org/blogs/blog-cacm/155468-what-does-big-data-mean/fulltext
Sun, S., Cegielski, C. G., Jia, L., & Hall, D. J. (2018). Understanding the Factors Affecting the Organizational Adoption of Big Data. *Journal of Computer Information Systems*, *58*(3), 193–203. https://doi.org/10.1080/08874417.2016.1222891
Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, *18*(7), 509–533.
The Economist. (2018). Europe’s tough new data-protection law. Retrieved December 6, 2018, from https://www.economist.com/business/2018/04/05/europes-tough-new-data-protection-law
Tippins, M. J., & Sohi, R. S. (2003). IT competency and firm performance: Is organizational learning a missing link? *Strategic Management Journal*, *24*(8), 745–761. https://doi.org/10.1002/smj.337
Turner, A. N. (1982). Consulting Is More Than Giving Advice. *Harvard Business Review*, (SEP).
Van der Veen, M., Van den Born, A., Smetsers, D., & Bosma, B. (2017). *Ondernemen met (big) data door het mkb*.
Verma, S., & Sekhar, B. S. (2017). Perceived strategic value-based adoption of Big Data Analytics in emerging economy: A qualitative approach for Indian firms. *Journal of Enterprise Information Management*, *30*(3), 354–382.
Wamba, S. F. (2017). Big data analytics and business process innovation. *Business Process Management Journal*, *23*(3), 470–476. https://doi.org/10.1108/09574090910954864
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 of Production Economics*, *165*, 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031
Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. fan, Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356–365. https://doi.org/10.1016/j.jbusres.2016.08.009
Wozniak, G. D. (1987). Human Capital , Information , and the Early Adoption of New Technology. *The Journal of Human Resources*, *22*(1), 101–112. https://doi.org/10.2307/145869
Wu, F., Yeniyurt, S., Kim, D., & Cavusgil, S. T. (2006). The impact of information technology on supply chain capabilities and firm performance: A resource-based view. *Industrial Marketing Management*, *35*(4), 493–504. https://doi.org/10.1016/j.indmarman.2005.05.003
Zahra, A. S., Sapienza, J. H., & Davidsson, P. (2006). Entrepreneurship and Dynamic Capabilities: A Review, Model and Research Agenda. *Journal of Management Studies*, *43*(4), 917–955. https://doi.org/10.1111/j.1467-6486.2006.00616.x

Appendix A

Factors	Category	Description	Frequency	Rank
Relative advantage	IC	The characteristics of big data are perceived as being better than those of the idea it supersedes (e.g., its unique role for innovation, competition, productivity, customer value creation and good business problems solution).	90	1
Human resources	OC	The firm's human resources are adequate for the task of adopting big data (e.g., analysts, data scientists, data science experts).	61	2
Technology resources	OC	The firm's technology resources are adequate for the task of adopting big data. (e.g., hardware, software, storage infrastructure).	54	3
Management support	OC	Managers are willing to allocate sufficient resources and encourage the initiative adoption of big data (e.g., top executives responsible for data management, CIOs' willingness to adopt big data).	52	4
Cost of adoption	IC	Possible expenses related to big data adoption. (e.g., the costs of using big data technology, large initial investment required to embrace big data adoption).	52	4
Security, privacy and ethics concerns in collecting data	EC	Data collection from individuals causes individuals' security, privacy concerns. (e.g., legal implications of collecting customers' private information).	49	6
Technology readiness	OC	Organization has sufficient internal IT expertise and technological infrastructure to adopt big data (e.g., IT knowledge and skills within the organization).	43	7
Trading partner readiness	EC	Adopt big data in order to follow partners and maintain the firm's internal balance with them. (e.g., readiness of suppliers in external collaborations).	28	8
Complexity	IC	The characteristics of big data are perceived as being difficult to understand and use (e.g. the difficulty of learning related knowledge for employees who will use big data applications).	27	9
Regulatory environment	EC	Governmental agencies encourage firms to adopt big data by providing related support (e.g., legal environment, industry regulation, data protection regulations).	25	10
Uncertainty/risk concern	EC	Concerns regarding potential unexpected consequences related to big data adoption (e.g., data security related risk, uncertain profitability for big data adoption).	24	11
Institutional based trust	EC	The firm's belief that it will be safe to adopt big data (e.g., reliability, reliable platform, system trust safeguard, "strong relationship of trust", inter-organizational trust).	24	11
Organization/IT structure	OC	Organization has a well-organized structure that is well-suited to the adoption of big data (e.g., cross-organizational collaboration structure, IT departments and staff configurations).	22	13
Decision-making culture	OC	Top managers' decision making at the firm level (e.g., culture of evidence-based decision making, decision-making norms).	19	14
Business strategy orientation	OC	An organization strategy that is oriented to business analytics and using big data for strategic decisions (e.g., business strategy of prioritizing big data adoption).	16	15
Business resources	OC	The firm has business resources that are adequate to the task of adopting big data (e.g., information sharing culture in the organization, policy toward big data initiative).	15	16
Change efficacy	OC	Organization members can easily handle the changes triggered by the adoption of big data (e.g., employees could handle the changes due to big data adoption with ease).	15	16
IS strategy orientation	OC	The firm's IS strategy prioritizes big data usage (e.g., information strategy, information governance policy).	13	18
Competitive pressure	EC	The extent of the pressure from a firm's competitors that can be combatted by the adoption of big data (e.g., competitive market, external threats from competitors).	13	18
Firm size	OC	The firm's annual revenue and number of employees that could support the adoption of big data (e.g., leading companies with more revenue).	11	20
Appropriateness	OC	The timing of the adoption of big data is advantageous for the organization (e.g., the organization potentially benefits from the introduction of big data if implemented at this time).	10	21
Compatibility	IC	The characteristics of big data are perceived as being consistent with the existing IT architecture in an organization (e.g., scalability, integration into the existing information systems).	10	21
Market turbulence	EC	Changes in customers' product preferences, demands, and needs in a big data environment (e.g., business environment fluctuation, changes in customers' product preferences).	9	23
Observability	IC	The characteristics of big data are perceived as being beneficial after observing how other organizations use it (e.g., a witness from the potential adopter observing a big data adoption by another firm).	1	24
Trialability	IC	The characteristics of big data are adopted without total commitment (e.g., it can be easily tried out with minimal investment).	1	24
IS fashion	EC	Information is obtained through external communication channels by focusing on an organization's peers and perceived experts such as vendors and customers (e.g., managerial interactions, social channels and trends).	1	24

Note: IC: Innovation Characteristics; OC: Organization Characteristics; EC: Environment Characteristics.

Figure 11 - Factors affecting the organizational adoption of big data (Sun et al., 2018)

Appendix B

Interview – Adoptie van big data technologieën door het mkb

Interviewer	Abel Bokdam
Project	Master sciptie Innovation Sciences, Technische Universiteit Eindhoven
Email	a.bokdam@student.tue.nl
Telefoon	

Dit interview is onderdeel van mijn master scriptie onderzoek. Ik onderzoek de adoptie van *big data* technologieën binnen het Nederlandse midden- en kleinbedrijf (mkb). Sommige bedrijven gebruiken al technologieën als *Hadoop, cloud computing* of visualisatie software om met grote datasets om te gaan. Ik wil echter onderzoeken wat de redenen zijn en wat er nodig is om deze technologieën succesvol te gebruiken binnen het mkb. Vanwege uw expertise en positie zou ik u graag vragen om mee te doen aan dit interview. Hieronder volgt meer informatie over het onderzoek en over uw rechten als deelnemer.

1. Het doel van dit onderzoek is om te begrijpen welke potentiële randvoorwaarden en barrières er bestaan voor de adoptie van big data technologieën door mkb in Nederland. Dit onderzoek draagt bij aan een beter begrip van het adoptie proces binnen mkb en hoe het mkb baat kan hebben van big data technologieën. Door voorwaarden en barrières te onderzoeken kan hier optimaal op ingespeeld worden door bedrijven en beleidsmakers.

2. Uw deelname is anoniem. Er zal een audio opname van dit interview gemaakt worden om te helpen bij het transcriberen. Nadat het transcriberen afgerond is, zal de audio opname vernietigd worden. Het transcript zal offline bewaard worden en data zal alleen gedeeld worden met de begeleider van mijn thesis. Nadat het onderzoek is afgerond, zal alle data vernietigd worden. U kunt uw deelname aan dit onderzoek op elk moment voor, tijdens of na het interview intrekken. In dat geval zal uw data vernietigd worden en niet worden gebruikt in het onderzoek.

3. Mocht u vragen hebben tijdens of na het interview kunt u deze altijd stellen. U kunt ook contact opnemen via mijn gegevens hierboven.

4. Door verder te gaan met het interview stemt u in met de hierboven genoemde voorwaarden.

Naam geïnterviewde	
Titel	
Organisatie	
Aantal werknemers	
Opgericht in	
Type bedrijf	
Contact informatie (email / telefoon)	
Datum van interview	
Locatie van interview	

Interview vragen

1. Kunt u wat over uw organisatie en uw rol vertellen?

2. Wat verstaat u onder big data en hoe wordt dit gebruikt in uw organisatie?
    a. Welke technologie en sinds wanneer? Waarom?
    b. Zijn mkb'ers klanten en waarom?
    c. Is big data gebruik projectbasis of langdurig?

3. Waarom heeft uw organisatie besloten big data technologieën aan te bieden?
    a. Hebben concurrenten invloed gehad op uw beslissing om (extra) big data technologieën te adopteren? Hoe?
    b. Hebben leveranciers invloed gehad op uw beslissing om (extra) big data technologieën te adopteren? Hoe?
    c. Hebben klanten invloed gehad op uw beslissing om (extra) big data technologieën te adopteren? Hoe?

4. Welke middelen zijn belangrijk in het mogelijk maken of verhinderen van de adoptie van big data technologieën in uw organisatie of bij de klant?
    a. Naast de middelen die u genoemd heeft, welke van de volgende factoren hebben invloed gehad op de adoptie van big data technologieën in uw organisatie?
        i. Expertise en vaardigheden van personeel
        ii. Technologische infrastructuur
        iii. Financiële middelen
        iv. Flexibiliteit om te veranderen

5. Wat voor acties heeft het management of de eigenaar van uw organisatie en die van de klant ondernomen om de adoptie van big data technologieën te faciliteren?
    a. Heeft uw organisatie ervaring in de adoptie van informatie technologieën?

6. Hadden werknemers een rol in de selectie van technologieën voor uw organisatie of die voor de klant?

7. Had u enige zorgen of bedenking met betrekking tot integriteit, veiligheid, privacy of ethiek voor of tijdens het gebruik van big data technologieën?

8. Hoe heeft beleid en/of regulering invloed gehad op de adoptie van big data technologieën in uw organisatie?

9. Wat zijn op dit moment de grootste barrières of succesfactoren om big data te gebruiken?
    a. Kan big data wel voordelen bieden en zo ja: wanneer dan? (bv. Eigen data, open data, type bedrijf b2c vs b2b)

Appendix C

Coding process example

Snippet of text	**Code**
Het zijn 500 miljoen events per dag bijvoorbeeld. Die krijgen we ingeladen op de Vodafone server, dus wij zien dat niet eens. Het zit voor ons in een database in een tabel en in principe zijn het gewoon tabellen. Je kan er een csv uithalen bijvoorbeeld. En als je het zou downloaden zou het wel meevallen, de hele database zou zo'n 2-3 terabyte zijn.	This was coded as **big data definition**, because the participant is talking about the size and velocity of the database they use.
We mogen niet binden aan gemeentelijke basis administratie, dus we kunnen ook niet toetsen of diegene nog in dezelfde plaats woont. Dat is gevoelig.	This was coded as **regulatory environment**, as this is about normative regulation
Grote organisaties... Dat zijn allemaal lijnclubs, die mogen helemaal geen geld uitgeven. Die worden allemaal in een lijn gestuurd en mogen geen programma's doen. Heel hiërarchisch. Iedereen heeft zijn potje en verantwoordelijkheid en als het morgen geen geld oplevert dan kunnen ze het niet kopen, want ze kunnen het niet verantwoorden. Dus dat moet eerst naar de directie en die moet zeggen, nou dat past wel in onze IT visie, dus dat gaan we proberen. Dus een bepaalde mate van flexibiliteit is wel belangrijk.	This was coded as **change efficacy** and **decision-making culture**, because the respondent is mentioning how within large organizations decisions are made in a hierarchical structure, which in turn affects the flexibility for change of an organization.

Appendix D

To increase the validity of the findings, a literature review was performed. These extra sources are part of the triangulation process as described in the methods chapter. When multiple sources converge to the same theme, validity is added to the findings of the research. This section will briefly discuss how the literature review was performed and provides an overview of the findings.

The starting point of the literature review was a search aimed at the adoption of big data and big data technologies. The idea is that articles studying this subject are likely to discuss certain challenges or issues faced with the adoption of big data. In turn, these challenges can be used to triangulate the findings of this research.

Table 11 on the next page presents the search terms per database, the number of hits and the articles that were selected. Articles with other units of adoption than SMEs, like organizations or firms in general were also selected. This is to compensate for the few academic articles on both SMEs and big data technologies.

The search queries were searched throughout topics and titles and across all years in the Web of Science and Scopus databases on the 2nd of June 2018. In the Google Scholar database the search query was limited to the time-period 2012 to the present, as it yielded too many irrelevant hits. This year was chosen as the popularity of the search term *big data* increased in 2012 (see Figure 13). Results were sorted on number of citations and the first 30 English articles were assessed on the relevance to this research. If the article was deemed relevant, it was added for analysis.

The articles are analyzed with the use of qualitative content analysis (Mayring, 2000). This is a method to qualitatively analyze text. Each article was read and its focus, research method, country, findings and adoption influencing factors were coded in a spreadsheet (see Table 12). This is an easy method to summarize articles and pair them with the factors influencing the adoption of big data. These findings were used to check the findings of the interviews of this research.

Table 11 - Search terms and selected articles for triangulation

Search terms	Database	Nr. Of hits	Author (year)
Big data Adoption SME	Web of Science	10	(Coleman et al., 2016)
	Scopus	4	(Coleman et al., 2016)
	Google Scholar^ 2012-Present	17.600	(Shah et al., 2017)
Big Data Adoption Firms	Web of Science	50	(Kwon et al., 2014) (Verma & Sekhar, 2017) (Sun et al., 2018)
	Scopus	58	(Kwon et al., 2014)
	Google Scholar^ 2012-Present	41.200	(Kwon et al., 2014) (Al-Qirim et al., 2017)
Big Data Adoption Small Firm	Web of Science	9	(Verma & Sekhar, 2017)
	Scopus	13	
	Google Scholar^ 2012-Present	18.900	
Big Data Analytics Adoption Firm	Web of Science	14	(Kwon et al., 2014) (Verma & Sekhar, 2017) (Sun et al., 2018) (Ramanathan et al., 2017)
	Scopus	16	(Kwon et al., 2014) (Verma & Sekhar, 2017) (Sun et al., 2018) (Caesarius & Hohenthal, 2018) (Agrawal, 2015)
	Google Scholar^ 2012-Present	17.800	

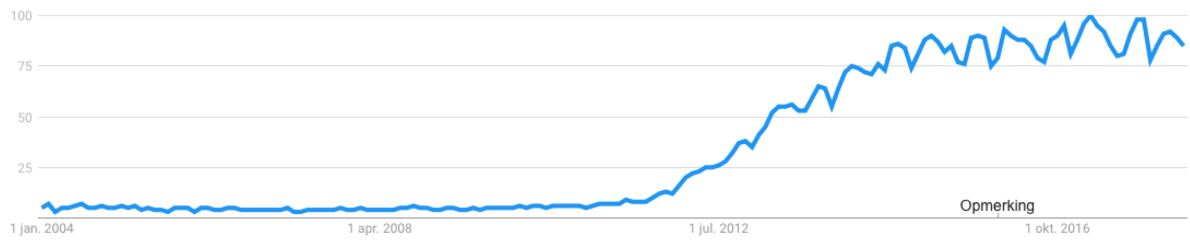

Figure 12 - Google search history of 'big data' from 2004 to May 2018.

The values in Figure 13 are relative to the peak ('100') popularity of the search term (Google Trends, 2018).

Table 12 - Analysis of articles for triangulation

Articles, Research Methods, Country	Focus	Findings
(Coleman et al., 2016), Qualitative, Literature Study, Europe	This paper identifies barriers to SME uptake of big data analytics and recognizes their complex challenge to all stakeholders, including national and international policy makers, IT, business management and data science communities.	The problems for the SME sector are complex, multifaceted and transversal over various dimensions like IT, data analytic intelligence, organizational structure, managerial models, capital structure and requirements, consulting, labor market, data security and legal aspects. These factors negatively impact big data adoption: - Lack of understanding - Dominance of domain specialists (reduced general awareness) - Cultural barriers and intrinsic conservatism - Shortage of in-house data analytic expertise - Bottlenecks in the labor market - Lack of business cases - Shortage of useful and affordable consulting and business analytics services - Non-transparent software market - Lack of intuitive software - Lack of management and organizational models - Concerns on data security - Concerns on data protection and data privacy - Different venture concept - Financial barriers
(Shah, Soriano, & Coutroubis, 2017), Qualitative, Literature Study, World	An investigative study on the concept of Big Data and its challenges towards implementation in manufacturing SMEs. Big Data aims to facilitate the collaborative approach in SMEs through the creation of real time data visualization to address key challenges to many of the market variations for every sector SMEs.	Companies have to get used to living in this information ecosystem to get more intuitive, anticipatory and personalized performance. Investment in data scientists, data warehouses and data analytics software is important to successfully adopt big data. A change in organizational interaction and learning is necessary, as well as expressing interest and beliefs towards accepting big data as a change tool for future growth.

(Kwon, Lee, & Shin, 2014), Quantitative, Survey, Korea	A research model is proposed to explain the acquisition intention of big data analytics mainly from the theoretical perspectives of data quality management and data usage experience.	A firm's intention for big data analytics can be positively affected by its competence in maintaining the quality of corporate data. Moreover, a firm's favorable experience (i.e., benefit perceptions) in utilizing external source data could encourage future acquisition of big data analytics. Surprisingly, a firm's favorable experience (i.e., benefit perceptions) in utilizing internal source data could hamper its adoption intention for big data analytics
(Verma & Sekhar, 2017), Qualitative, Interviews, India	To provide an insight about factors affecting Big Data Analytics (BDA) utilization and adoption in Indian firms.	The major reason behind BDA non-adoption is that the organizations did not realize the strategic value of BDA, and they were not ready to make the changes because of technological, organizational and environmental difficulties.
(Sun, Cegielski, Jia, & Hall, 2018), Qualitative, Literature Study, World	This study applies the results of a content analysis to develop a framework to identify the main factors affecting the organizational adoption of big data.	The results indicate that the innovation characteristics consist of relative advantage, cost of adoption, complexity, compatibility, trialability, and observability. The organization characteristics encompass human resource, management support, technology resources, technology readiness, decision-making culture, change efficiency, business strategy orientation, IT/organization structure, business resources, IS strategy orientation, firm size, and appropriateness. Security, privacy and ethical concerns in collecting data, trading partner readiness, regulatory environment, IS fashion, market turbulence, and institutional based trust represent the environment characteristics.
(Al-Qirim, Tarhini, & Rouibah, 2017), Qualitative, Literature Study, World	To develop an adoption model of Big Data that could detect key success predictors.	The research finds a great interest and optimism about BD value that fueled this current buzz behind this novel phenomenon. Like any disruptive innovation, its assimilation in organizations is oppressed with many challenges at various contextual levels. BD would provide different advantages to organizations that would seriously consider all its perspectives alongside its lifecycle in the pre-adoption or adoption or implementation phases. These facilitating factors are: - Organizational trust - Big Data analysts and decision-makers - Big Data domain capabilities

		- Knowledge exchange
(Ramanathan, Philpott, Duan, & Cao, 2017), Qualitative, Case Study, UK	Understanding issues faced by retail firms when they start a project of implementing business analytics (BA) and understanding the impact of BA implementation on business performance	There appears a link between adoption of BA and business performance (including performance in terms of environmental sustainability), and this link is moderated by the level of BA adoption, IT integration and trust.
(Caesarius & Hohenthal, 2018), Qualitative, Case Study, Scandinavia	Little attention has been paid to the challenges that many incumbent organizations face when they try to explore a possible adoption of big data technologies. This study investigates how incumbents handle such an exploration and what challenges they face.	Big data is to some extent different in two aspects: the concept is elusive and can mean different things to different firms, and it can have a transformative effect on the organization of work in the firm. This elusiveness makes it difficult to explain what the technology means but also opens up possibilities of defining the technology in creative ways, thus making it possible to gain support and funds to introduce it. The transformative capability of big data makes managers wary as it might threaten their position in the firm, and creates ripple effects, transforming other systems besides those directly connected to the technology.
(Agrawal, 2015), Quantitative, Survey, China & India	Investigates the determinants that influence BDA adoption in context of the firms from two big emerging economies of Asia –China and India.	Six variables i.e., complexity, compatibility, regulatory support, organizational size, competition intensity, and environmental uncertainty were found to be significant determinants of BDA adoption, and three variables i.e., relative advantage, absorptive capacity, and technological resource competence were found to be non-significant determinants. Out of the six determinants, regulatory support and complexity are inhibitors, and other determinants are facilitators of BDA adoption.