UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Challenges of implementing Big Data in large organisations

- *A case study*

Authors: Linda Carlsson
Da Yu

Supervisor: Henrik Agndal

Acknowledgement

University of Gothenburg School of Business, Economics, and Law
Gothenburg, 2018-06-03

Linda Carlsson   Da Yu

Abstract

This study examines the challenges faced by large organisations when implementing Big Data to support the decision process. Existing literature on Big Data implementation are limited and have not been focused on how the challenges affects the decisions at an organisation. However, previous research has focused on the implementation challenges of IT tools, as well as studies on how Big Data can be used to support decisions. In order to expand upon existing knowledge, this case study is focused on Big Data implementation challenges and its implications on the decision process. The study is conducted as a case study at Volvo Car Group (VCG), with the main interviewees being data scientists, business employees, and managers, from diverse departments. The interviews were eight in total, focusing on how Big Data is being used, how it can be used, how it should be used and the challenges that are preventing that from occurring. Despite the observed challenges, the interviewees confirm that even if there are some challenges regarding Big Data, there are a large potential of value creation for the organisation, if used correctly.

The study identifies several challenges faced by the Big Data implementation, these are categorized into *technical*, and *managerial* challenges. The main technical challenges are: *poor data quality, data restriction, data silos*, and *an inefficient Big Data process*, all influencing the usage and implementation of Big Data negatively. The main managerial challenges are: *inexperienced staff, limited training, communication*, and *poor teamwork* between the business and IT departments.

The result of the study indicates that there are several aspects that must be taken into consideration, in order to achieve a successful implementation of Big Data, to support the decision process. The most important aspect is to have a good communication between all affected parties, since good communication and understanding of the challenges will provide less friction between those involved, as well as being able to contribute to a somewhat general definition of Big Data.

The recommendations from the study is that VCG should focus on solving the challenges that lowers the demand for Big Data analysis - mainly to strengthen their data focused culture as well as improving their Big Data process in order to produce better and faster results. This should be done through starting a new department or work group where a part the business side and IT side is combined to focus on the Big Data implementation challenges while improving the business-IT relationship at the same time.

Key words used: *Big Data, Decision Making, Decision Support, Information Systems, Implementation Challenges.*

Table of Content

1. Introduction

The introduction chapter consist of a presentation of Big Data in general, Big Data implementation challenges as well as how Big Data can support decision making. In addition to this, the chapter will illustrate some of the gaps in the current research, this will then culminate into the purpose of the study. Concluding this chapter is the disposition of the thesis.

The aim of this thesis is to obtain insights concerning key challenges of implementing Big Data for decision-making support within an organisation. The aim has been chosen due to the fact that, management accounting (and thus the decision-making process) has much to gain from implementing business intelligence (BI) and analytics for managerial accounting tasks. It also allows the management accountants to participate and have a more active role regarding data creation and decision support (Rikhardsson & Yigitbasioglu, 2018). In addition to this, Big Data has become increasingly popular, due to the continuous digitalisation of the world, creating more available information (Manyika et al. 2011) hence IT becoming an essential carrier of accounting information (Granlund, 2011). There is therefore a need to investigate the theories and methodologies applied in the accounting information systems field, in order to be able to understand the complexity and relationship for new technologies (Granlund, 2011).

> *"Many organisations are therefore implementing business intelligence & analytics (BI&A) technologies to support reporting and decision-making" (Rikhardsson & Yigitbasioglu, 2018)*

Big Data is a rather new concept within IT, referring to the dynamic, large and diverse volumes of data being created by people, tools, and machines (Davenport & Dyche, 2013; Vloet, 2016). Warren, et al. (2015) state that Big Data can offer an unprecedented level of potential, providing diverse, large datasets and sophisticated analyse. This allows a definition of how data are accumulated, recorded and most importantly, how data can be used for effective and efficient decision support in order to achieve organisational goals (Warren, et al 2015).

Moreover, it has been noted that in accounting information research the production of information for management control and decision making, has not been its main focus. Surprisingly, since one of the main functions of information systems (IS) is to provide managerially relevant information (Granlund, 2011). The current focus of analytic research has highlighted it as being perceived as contributing with importance to the management of an organisation, for example in decision making (Rikhardsson & Yigitbasioglu, 2018).

The concept of Big Data is, however, still considered difficult to understand and many companies lack experience in how to use Big Data within the organisation. The term of Big Data is, by most, connected to the size of the data, there are, however, more characteristics defining the concept of Big Data (Anon, 2017). How to approach the emerging subject of Big Data is therefore highly important, since the decisions made today, will affect the future (Boyd &

Crawford, 2012). The increasing stream of Big Data requires organisations to consider new ways of making decisions with the usage of this new source (Davenport, 2014).

The global competition among organisations creates a need for better decision- and control-support, directly affecting the value chain. IT is therefore a very important factor, since it directly affects the value of an organisation (Granlund 2011).

"Traditionally, management accounting is the primary support for decision-making and control in an organisation, As such, it has clear links to and can benefit from applying BI&A technologies … little research has focused on this link" (Rikhardsson & Yigitbasioglu, 2018)

Even though there are opportunities for management accounting and decision making when studying business intelligence, there are still no largely spread understanding of this within the accounting academia (Rikhardsson & Yigitbasioglu, 2018). McAfee and Brynjolfsson (2012) note that there are challenges in order for an organisation to implement the usage of Big Data. These challenges may be of both technological and/or managerial in its appearance (Vloet, 2016). Technological challenges may be characterized by limitation of the IT infrastructure, privacy and security issues. McAfee and Brynjolfsson (2012) mention, in addition to the technical challenges, the managerial challenges. In order to implement Big Data efficiently, good professionals and organisational culture are highly important to take into consideration (McAfee & Brynjolfsson, 2012). Moreover, according to the Global data management benchmark report (2017)

"While most organization around the global say that data supports their business objectives, less than half of the organization globally (44%) trust their data to make important business decisions."

Indicating that the lack of trust for the data can be a significant challenge, when implementing Big Data. The need for standardised systems as well as disciplined data policies on how to implement Big Data is therefore highly important, otherwise the organisation may miss opportunities, suffer decreases in brand value, or customer satisfaction (Global data management benchmark report, 2017; *Interviewee 2,* 2018). This is also confirmed by Ahmed, V. et al., (2017), who state that in the architecture, engineering and construction industry, a gap and imbalance has been identified between the data analytics in use and the data capture.

"While it seems to be widely acknowledged that IT plays an important role (and increasingly so) in the field of accounting, the relationships between IT and accounting, especially as regards management accounting and control, has been studied relatively little" (Granlund, 2011)

The managerial and technological challenges faced by an organisation when implementing Big Data, makes the possibility to efficiently implement Big Data questionable. Furthermore, Frizzo-Barker et al. (2016) argue, that since Big Data is a new, emerging concept, the research of Big Data is not widely spread nor developed. Most importantly, existing researches on Big Data is

primarily focused on theorisation and formulation of expectations, and often emphasises on the positive aspects of Big Data (Vloet, 2016). The connection between IT (Big Data) and management control is therefore highly important to study, since knowledge is power, and IT (including Big Data), creates new important knowledge for the organisation and its management. Even though this has been stated, there is still a lack of understanding of the interplay between the technology and management accounting and control (Granlund, 2011).

However, the awareness of the challenges regarding Big Data has been raised in recent years, but the existing literature has mainly paid limited attention challenges towards the use of Big Data (Vloet, 2016). Indicating a need to identify and examine the challenges connected to the implementation of Big Data. By focusing on the challenges of implementing the usage of Big Data in organisations, new insights regarding different problems, could be gained by the managers (Tesfaye, 2017). Moreover, is the existing research within the field, according to Rikhardsson and Yigitbasioglu (2018), based upon conceptual rather than empirical research. Indicating that this study, is among the first case studies regarding Big Data, its challenges and its effect on decision making in practice.

Due to the limited research, and thereby the limited understanding of the outcomes of the implementation of IT focusing Big Data in an organisation, makes this highly important to examine. This since Big Data and IT contribute with forming the management control, and decision making within an organisation. How to control and make decisions can therefore not be studied without the technology, since the technology and what is included within, affects the control and the decisions made (Granlund, 2011).

1.1 Purpose

The purpose of this thesis is to identify and examine challenges occurring when implementing Big Data for decision making purposes. In this study the definition of *"implementation of Big Data"* is regarded as an ongoing process, which might require several years to fully implement (Mabert, et al., 2003; Howson, 2013).

1.2 Disposition

The disposition and the content of this thesis in illustrated in, *figure 1:*

1. Introduction	• Contains background information, including problimisation, and gap in current research regarding Big Data Challenges
2. Theretical Framework	• Explains important concepts, information, and theories and is used as a foundation for identifying and examine Big Data challenge
3. Methodology	• Explains how the thesis was conducted in order to answer the research questions, as well as why a case study was chosen
4. Empirical Findings	• Contains the information collected from the conducted interviews
5. Analysis	• Analysis of the empirical findings, focusing on the current Big Data implementation challenges for the decision process
6. Discussion	• Presents the results from the analysis of the findings compared and discussed toghether with the information from the Theoretical Framework
7. Conclusion	• The conclusion from the analysis and discussion is presented, togheter with recomendations for practice, and ideas for future research topics

Figure 1: Deposition

2. Theoretical Framework

The Theoretical Framework will present the concept of Big Data and other IT tools, comparing the similarities and differences between them in order to gain a deeper understanding of the different concepts. This will contribute to the formulation of the research question(s). The section will be concluded with a summary of the Theoretical Framework and an explanation of its purpose.

2.1   Big Data
2.1.1   Definition of Big Data

The general conception of Big Data is that it has the potential to change and improve the creation of business value (Rikhardsson & Yigitbasioglu, 2018). Big Data is a concept containing a wide variety of data that can be used in business intelligence (Howson, 2013). BI can facilitate the organisation to understand general trends, strategies of their competitors, as well as the environment of the organisations operations in order to make decisions (Negash, 2004). Howson (2013) defines BI is *"a set of technologies and processes that allow people at all levels of an organization to access and analyze data"* as well as the culture and creativity of the employees in order to use the data. Big Data however, cannot according to Manyika, et al (2011) be used in traditional BI systems since:

> *"Big Data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage and analyze"* (Manyika, et al., 2011)

Howson (2013) states in similarity to Russom (2011) that Big Data has a clear definition except the size of the data. This definition is a constitution of *"the 3 V's of Big Data"* which is *Volume*, *Velocity* and *Variety*:

- **Volume**: According to Howson (2013), the data used in traditional BI is counted in gigabytes while Big Data is counted in the range of petabytes (one petabytes equal one million gigabytes). The scope of Big Data varies widely and is therefore difficult to quantify (Howson, 2013; Russom, 2011).
- **Velocity**: Both Howson (2013) and Russom (2011) identifies the velocity of new incoming data (for example streaming data) as a characteristic of Big Data, and the pace required for making decisions differs from the use of traditional data in BI, compared to Big Data. (Howson, 2013; Russom, 2011).
- **Variety**: In Big Data, data from various sources are included, being one of the most prominent characteristics of Big Data (Howson, 2013). Different types of data included in the analytics of Big Data are for example unstructured data (text and human language), semi-structured data (XML, RS feeds), and structured data, as well as data that is hard to categorize such as audio and video (Rossum, 2011).

The information regarding a clear definition of Big Data is to some extent contradictory to other authors, stating that the term Big Data lacks a formal definition and it has been used in several inconsistent meanings (De Mauro, et al., 2016). The term has previously been used to refer to large data sets requiring super computers, but as the technology moves forward, what once required these super computers can today be analysed on ordinary computers (Boyd & Crawford, 2012). According to Davenport & Dyché (2013), is Big Data not about the volume, but rather the variety of data. The most important aspect of Big Data is the possibility to analyse vastly diverse data, not managing large data sets (Davenport & Dyché, 2013). The diverse data and its lack of structure is the most difficult aspect of Big Data, and not the volume of the data, presenting both new challenges and new opportunities (Davenport, 2014). These opportunities are for example linked to the usage of more sources and types of available and usable data, contributing with new information and insights to create predictions and decisions. Examples of new sources to be used in management accounting and decision processes are for example video, audio, and textual data. The technology, allowing real-time access to this information, is highly likely to affect the decision-making process for the managers (Rikhardsson & Yigitbasioglu, 2018).

The definition of Big Data and its volume is therefore relative, and dependable upon diverse factors, such as: time, type of data, and the industry using the data (Gandomi & Haider, 2015). One of the reasons for the lack of formal definition is the fast and chaotic development of Big Data during the recent years (De Mauro, et al., 2016). This can be illustrated by the *figure 2* below, illustrating themes and related topics concerning Big Data:
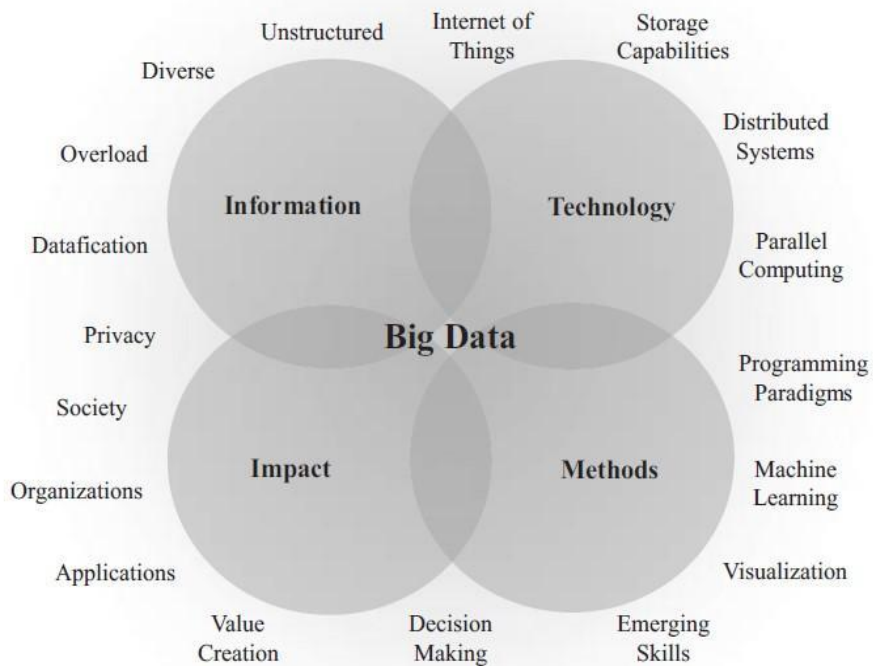
Figure 2. (De Mauro, A., Greco, M. & Grimaldi, M., 2016)

De Mauro, et al, (2016) do however agree that the fundamental characteristics of Big Data are the three Vs; Volume, Velocity, and Variety, requiring specific technology, and analytic tools in order to transform the information from Big Data into something valuable (De Mauro, et al., 2016) in accordance with Ward and Barker (2013), who states that the existing definitions of Big Data, have some similarities, concerning the size of the data sets (volume), the structure and behaviour of the datasets (velocity and variety), and in addition to the three V:s is the tools and techniques required when using the datasets (technology) (Ward, & Barker, 2013).

De Mauro, et al, (2016), and Ward, and Barker (2013) states that the different definitions of Big Data is dependable upon different stakeholders and their perspectives. De Mauro, et al (2016) continues with stating that they have found four distinct groups in which the different definitions of Big Data can be divided: Attributes of Data, Technology requirements, Overcoming thresholds, and Social Impact. (De Mauro, et al., 2016).

Gandomi & Haider, 2015 has in their paper stated some additional defining concepts regarding how organisations define Big Data:
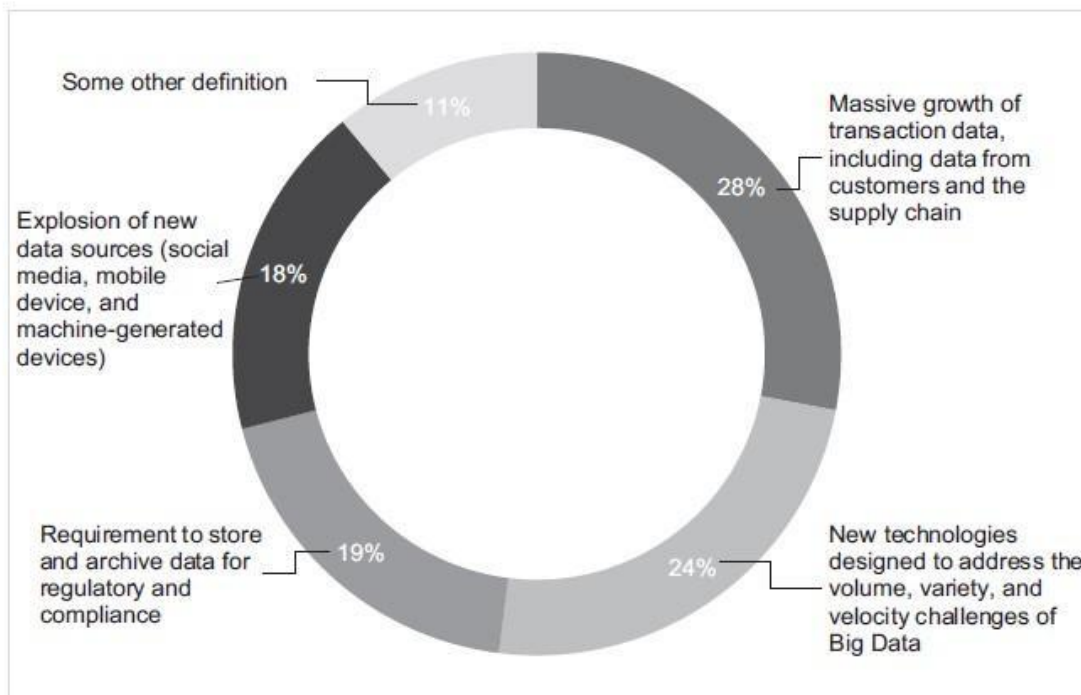
Fig. 2. Definitions of big data based on an online survey of 154 global executives in April 2012.

Figure 3. (Gandomi & Haider, 2015)

Another definition of Big Data is made by Boyd & Crawford (2012) who have stated their definition of Big Data as: "*A cultural, technological, and scholarly phenomenon*". The definition is dependent on the *technology*, and its ability to analyse the data correctly, the *analysis*, and its

possibility to find correlations among the data sets, and, the *Mythology* surrounding Big Data, especially the belief that large volumes of data automatically creates a deeper and better understanding, especially in areas where it was previously impossible to do so (Boyd & Crawford, 2012).

2.1.2   Big Data vs. Small Data

According to Davenport (2014) is the main purpose of the more known concept of "*small data*", to support internal decisions, such as what to offer the customers, and pricing. Big Data on the other hand, can offer in addition to this, new dimensions such as discovering new business opportunities, meaning that instead of solely producing advisory reports to the management, Big Data analytics contributes with customer-facing products and services (Davenport, 2014).

Big Data focuses on large volumes of unstructured, fast-moving data, creating the need for new management approaches, regarding the usage of Big Data for internal purposes. This is due to the fact that the large volumes of data are continuously increasing and flowing (Davenport, 2014). Big Data does not solely refer to the size of the data, but rather about the capacity to search aggregate, and cross-reference large data sets (Boyd & Crawford, 2012). In traditional small data analyses, the data is extracted, put aside, analysed, to later form a decision model, when working with Big Data, this is not possible, instead it is necessary to have a continuous approach regarding the sampling, extraction, and analysing of the data, due to the fast-flowing data (Davenport, 2014).

Ahmed, et al. (2017) also contributes with explaining the distinctions between small and Big Data, which together with the definitions provided by Davenport (2014) can be seen summarised in *Table 1*:

	Small Data	Big Data
Memory	Centralised storage	Distributed sources
Architecture	Serial, centralised	Parallel, decentralised
Data types	Homogenous, static	Heterogenous, dynamic and evolving
Data management	Fits into a relational database and/or warehouse	Diverse formats may need to be integrated
Data quality	Well documented, strong correction techniques	May be unclear, large amount of uncertainty
Data processing	All the data has some kind of utility	Data must be processed continuously, may be difficult to find the useful data among all the data
Result analysis	Statistical, answers to specific questions	Non-statistically significant results may appear significant due to the size of the data, explorative discoveries

Table 1. (Ahmed, V. et al., 2017; Davenport, 2014)

2.1.3   Big Data Management

Big Data management is a concept that encompasses the policies, procedures and technology used for the collection, storage, governance, organisation, administration and delivery of substantial amounts of data (IDG, 2016). Big Data management is currently an actuality for an increasing number of organisations in various areas and represents a set of challenges involving Big Data modelling, cleansing, migration, analysis and visualisation (Rossi & Hirama, 2015).

According to Rossi and Hirama (2015) are the technological resources in addition to the employees and processes crucial aspects for all organisations in order to facilitate the management of Big Data. Big Data management can therefore be supported by the technology, employees, and processes. In addition, the authors highlight the human aspect as one of the most significant resources in order to sustain Big Data management (Rossi and Hirama, 2015). Manyika, et al (2011) also argue one major challenge for organisations to be the lack of analytic and technology skilled employees, who can make use of the Big Data. Acknowledging the most significant barrier to overcome being how to establish a data-driven organisational culture and structure (Manyika et al. 2011). Parmar, et al (2014) further discuss that by having a data-driven culture, the organisation recognises the data and analytics as a central function, and by promoting a cross-functional distribution of data allows more fact-based decisions (Parmar et al., 2014).

The usage of both Big Data and better analytical tools, will allow organisations to better evaluate the performance of the organisation, as well as its employees (Rikhardsson & Yigitbasioglu, 2018).

2.1.4   Challenges of Big Data management

There are different factors making the Big Data management more challenging, than managing smaller repositories of data, these challenges are of both of technological and managerial characteristics (George et al, 2014). According to Morabito (2015), is one of the main challenges, the management's lack of understanding of the potential contribution of value Big Data can induce to the organisation (Morabito, 2015). Manyika, et al (2011) formulates in their research:

"Organizational leaders need to understand that Big Data can unlock value-and how to use it to that effect" (Manyika et al., 2011).

Executive support is a fundamental aspect in order to succeed with the BI program at the organisation. In companies who describe their BI projects as having a significant impact, 92 percent have executive sponsorship, whereas in the projects described as having no impact, only 75 percent had executive sponsorship. The chief officer (CEO) as the BI sponsor, has the highest rate of BI success (Howson, 2013). However, a lack of business sponsorship is still the most relevant threats for Big Data management (Russom, 2013).

Another large challenge, complicating the Big Data management, according to Howson (2013), is the lack of a trained "hybrid business-IT person" who can act as a powerful bridge between the different BI stakeholders:

"The business derives the value and IT enables the systems" (Howson, 2013).

The partnership and trust between IT employees and business employees is essential to any of Big Data project. However, the business employees and IT employees have different mind-sets and speak two different technical languages. This causes frustration and friction between the business and IT (Howson, 2103).

Another problem with the of Big Data analytics, is the lack of coordination between the database systems. The analysts using Big Data has therefore to perform time-consuming processes, of first having to export the data from the database, then performing a non SQL process (such as data mining and statistical analyses), and then lastly bring the data back (Labrinidis, & Jagadish, 2012). But it is also important to remember that, since numbers do not themselves contribute with anything (Boyd, & Crawford, 2012), that:

"Getting the most from the data requires interpreting them in light of all the relevant prior knowledge" (Marx, 2013)

2.1.5 Big Data Implementation

Data analytics is more challenging then only to locate, identify, and understand the data. The collected data is not always in a possible format to analyse. The required information from the data has to be extracted, through an extraction process, and then transformed into a structured and suitable form in order to conduct the analysis. In addition to this, another prominent challenge regarding the data is incompleteness, scale, timeliness, and process complexity. For large-scale analyses, the extraction process must be automated, requiring the differences in data structure and semantics to be expressed in a computer understandable form in order for the data to be possible to analyse (Labrinidis, & Jagadish, 2012)

The results from the analyses must be interpreted, in order for the Big Data analyse to be useful, i.e. the users have to be able to understand the data. The interpretation of the analysis usually involves taking into account all the assumptions made prior to the data collection. This can be summarised into a multi-step schematic requiring the value to be extracted from the data. (Labrinidis, & Jagadish, 2012).

In order to be able to use and implement Data analytics, there are some challenges that must be taken into consideration. These aspects can be divided into two main areas: *Technical issues, and managerial challenges* (Ahmed, et al., 2017). Technical challenges are for example: *Data security and confidentiality, Inexperienced staff, Lack of business leadership, Difficulties in designing a analytic system, Data quality; the belief that a larger data set automatically becomes a good, Data presentation, and Lack of database software* (Ahmed, V. et al., 2017).

The managerial challenges differ from the technical challenges, and can for example be: *Difficulties to create leadership with clear goals, Difficult to ask the right questions, Finding competent employees to perform the analysis of Big Data, and The required technology is often too complex for the ordinary employees at the IT department* (McAfee et al., 2012).

Ahmed, et al., (2017) also states in addition to the technical and managerial challenges, some general challenges affecting the implementation of Big Data within an organisation. These challenges are for example: *Ethical concerns regarding the usage of the Data and its contents, and the accuracy of the sources* (Ahmed, et al., 2017)

2.2 Analogies between IT implementation, and Big Data implementation

As previously mentioned, there has been a limited amount of research conducted regarding the aspect of challenges of Big Data implementation (Vloet, 2016). Therefore, since Big Data is a new technology in IT (Davenport & Dyche, 2013; Vloet, 2016), it can be useful to study IT implementation challenges, as a basis for understanding the Big Data implementation challenges. In the following sub-chapters, more specific implementation challenges regarding different IT systems are presented as well as the presentation of commonalities and differences of challenges across type of IT systems.

2.2.1   Information technology and implementation

Information Technology (IT) refers to any form of technology, including any equipment or technique, used by people to handle information (Butterfield, 2016). More specifically, IT is the utilization of computers in order to store, retrieve, transmit and manipulate data (Daintith & Wright, 2008).

"BI&A (business intelligence & analytics) provides data-centric decision support to management accountants in, for example, planning, performance measurements and cost management techniques" (Rikhardsson & Yigitbasioglu, 2018)

By implementing analytics techniques to support management accounting and decision making, different challenges affecting the implementation process may occur (Rikhardsson & Yigitbasioglu, 2018). However, the implementation of IT systems can indicate and create a change in the organisation's existing technological architecture (Doyle, 2013). Therefore, is it common for IT implementation processes to encounter challenges (Winter, 2011).

Many IT experts estimate that more than 50% of all IT projects fail to meet their goals due to the fact that new IT systems take years of struggle to implement (Kehob, 2006). There are many reasons for the implications occurring at the implementation of IT systems, which is among the most challenging high-risk ventures any organisation will undertake (Doyle, 2013). The main reasons for the implications are (Kehob, 2006; Doley, 2013):

1) Lack of aligned key leaders.
2) Employees do not like the major change in new IT systems.
3) Accountability has not been created in order to take care of new IT systems (Kehob, 2006; Doley, 2013).

2.2.2   Implementation of ERPS

Enterprise resource planning systems (ERPS) are aimed for the integration of all corporate information into one central database, allowing all of the data being accessible from various resources through the firm (Dechow & Mouritsen, 2005). It directly influences the organisations processes and the execution of control (Tesfaye, 2017). ERPS affects the control structure of a organization and therefore significantly impacts management control (Quattrone & Hopper, 2004).

According to Chen and Lin (2009) is an ERPS a compound network consisting of different business processes. Although the ERP applications have been widely used by various industries, the challenges faced during and after implementation remain a growing concern (Momoh et al., 2008). Due to the complicated implementation processes including the fact that it requires more time and money than predicted, has resulted in the ERP industry not performing as formerly expected, since there is an inadequate understanding of how to implement (Momoh, et al., 2010).

The top ten critical challenges of implementing ERPS are, according to Momoh, et al. (2010), based upon 52 studies about ERPS implementation between 1997-2009. In addition to this, the authors highlight that these challenges mainly apply to large organisations. *Figure 4* illustrates, that during the implementing processes a lack of change management, poor data quality as well as poor understanding of ERPS business implications and requirements are the significant challenges (Momoh et al., 2010).
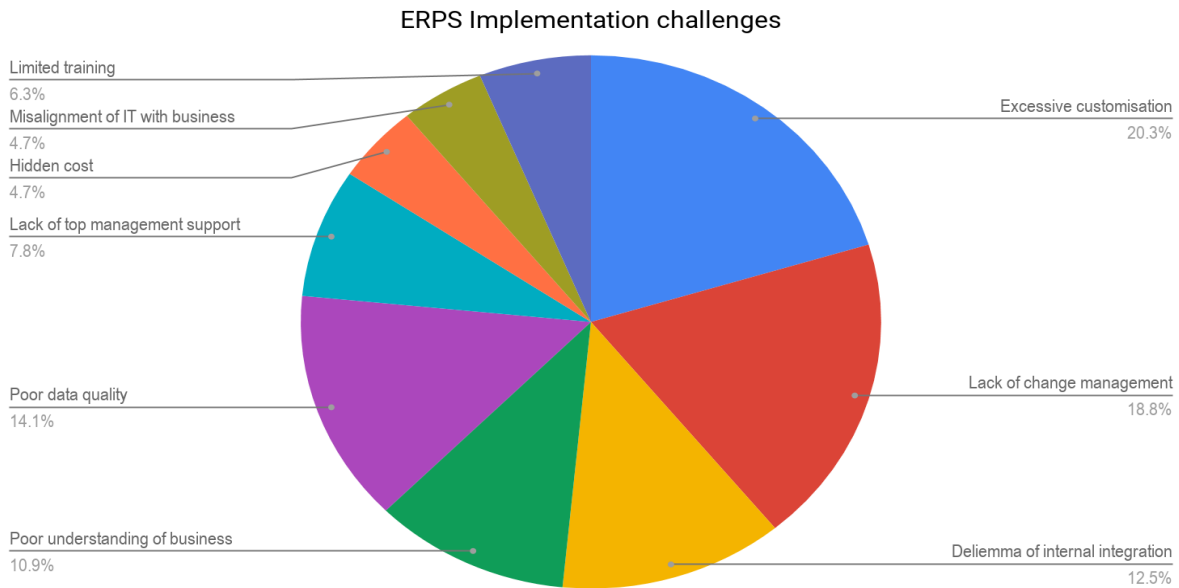
Figure 4: ERPS implementation challenges according to Momoh et al., (2010).

Ehie & Madsen (2005) reports that a lack of change management is one of the key organisational issues resulting in the failure of implementing ERPS. There are several factors included in ERP implementation with regards to the lack of change management (Al-Mashari, 2003). The lack of top management support, suitable people joining the implementation teams, and strong involvement of people from the field are main challenges that contributes to the resistance towards change in the ERP system implementation (Cissna, 1998).

During the ERPS implementing processes, poor data quality at operational level increases operational cost since resources are spent detecting and correcting errors (Momoh et al., 2010). Alshawi et al. (2004) argue that, the quality and accuracy of the data is a problem if the data goes into a system that is not being accurate or immediately accessible, which results in the whole ERPS becoming resented as a result. Due to the fact that ERPS integrate different information from each department of the organisation, poor data quality affects all operations (Momoh et al., 2008).

Another challenge regarding ERPS, is the attempt to implement ERPS before realising the full business benefits and the implementation barriers concerning the system (Momoh et al., 2010). Hence, one significant challenge in the ERPS implementation is the lack of understanding of the ERPS business implications and requirements (Ehie & Madsen, 2005). Another explanation is the existing gaps between the ERPS generic functionality and the organisational requirements (Soh et al., 2000). Therefore, without reconciling the technological imperatives of ERPS and the business needs of the organisation, the correct amount of resources needed for implementing the ERP solution cannot be allocated by the organisation (Kogetsidis et al., 2008).

2.2.3   Implementation of CRMS

CRMS (Customer relationship management system) is an information system tracking customers' interaction with the organisation. The system stores customers' information allowing employees to instantly and constantly extract data about different consumers (Nguyen et al., 2007). This contributes to the achievement of different business goals, for example attracting and retaining loyal customers (Swift, 2001). A successfully implemented CRMS can, according to Nguyen, et al., (2007), in addition to the achievement of business goals, also contribute with great business benefits to the organisation.

However, while CRMS could contribute with these benefits, many companies fail to implement the system (Nguyen et al., 2007). One of the most significant technical challenges encountered by CRMS is the storage of data in data silos, not cooperating with other databases. The information available in CRMS is therefore difficult to extract and share with other enterprise systems. The design of CRMS therefore becomes a silo, preventing the information from being possible to be exchanged with other applications, such as customer service, analysing and forecasting systems (Khamees, 2013). In addition, the technical skills needed for the system can also be a challenge, due to the (sales) employees having a limited position to synthesise information effectively when the CRM applications are utilised. This results in companies with mainly sales employees, who possess lower levels of technological expertise, will be less adept at providing richer and more highly processes information to managers (Ahearne et al., 2012).

Besides the technical issue, management has to consider and research the possible challenges which might be encountered during the implementation. Some of the main managerial challenges are the following which are cited by Ramsey (2003) and explained by Nguyen et al., (2007):

- *Lack of definition.* Many organisations are uncertain of where to start and what to abandon or acquire in this new type of information systems (IS) since CRMS is surrounded with new concepts. Also, management is not sure how to approach CRMS and how it would affect other parts of the organization's operations.

- *Poor leadership.* Functional heads have always been leaders of CRMS implementation projects. They often do not have enough strategic plans or perspective experiences of

CRMS. This is because, functional heads mostly focus on improving activities which related to their own functions, instead of working on the alignment of strategies toward to the whole organisation's goal. Therefore, lack of an authority leader who can align different internal functions on the same goal and mission leads to the failure of implementing CRMS.

Moreover, Zimmer (2006) finds that more than 50% of the companies who have implemented CRMS also met challenges after the implementation. The researcher highlights two principal reasons why CRMS does not always meet origin expectations. They are:

1) *The disconnection of CRM vision and execution* (Zimmer, 2006)*;*
2) *The rising standard of CRM excellence* (Zimmer, 2006)*.*

The first challenge contains two factors, one of them is the lack of research conducted prior to the implementation and having an appropriate plan before the implementation process began. The other factor is based on the fact that the project executions often fail or suffered from a lack of senior management support, poor project management, or poor training of users, who, when the system is up and running, do not know how to use it and/or maintain the system (Nguyen, 2007). Kovacs (2006) writes about the challenge of *"The rising standard for CRM excellence"*, due to the increasing competition and demand from customers. Meaning, that organisations must, in order to survive, outdo each other. Therefore, if organisations are rushing into the implementation of CRMS before the difficulties are known and recognised, the implementation will fail (Nguyen, 2007).

2.2.4 Summarized challenges applicable to Big Data implementation

Both ERP and CRM are enterprise systems, sharing commonalities regarding implementation challenges (Hendricks et al., 2007). The ERPS challenges can, according to, Themistocleous et al. (2001), be divided into two categories: *managerial*, and *technical* challenges. The challenges encountered during the process of implementing CRMS can be categorised alike (Nguyen et al., 2007). In different articles, the authors have emphasised on managerial rather than technical issues for both implementation processes. This is due to the fact that, in addition to developing the technical aspects of information systems, more effort is required in understanding the more complex management issues involved (Huang et al, 2003; 2004). Moreover, both ERPS and CRMS encounter challenges not only during the implementation process but also after it (Momoh et al., 2008; Zimmer, 2006).

Common challenges for both systems highlighted from the previous chapters are, a lack of understanding of how to implement the systems in order to support the needs of the organisation, and lack of management support. Langenwalter (2000) explains that this is due to organisations rushing into the implementation, without understanding the business implications. This leads to a gap between the functionality and the specific organisational requirements (Davenport, 1998).

The support and commitment from the management is highly important in order to have a successful implementation (Anon, 2006), since the lack of management support affects the systems greatly.

Besides commonalities, ERPS and CRMS also encounter implementation challenges from different perspectives. The main reason behind the differences, is due the core functionalities and purposes of the two enterprise systems being different. ERP focuses on reducing overheads and cutting costs, whereas CRM works to increase profits by producing greater sales volume (Hendricks et al., 2007).

2.3   Summary of the Theoretical Framework

The concepts presented in the theoretical framework provide fundamental concepts in order to be able to understand what Big Data is, and what challenges might arise when implementing Big Data analytics in an organisation. Due to Big Data being a relatively new concept, and being based on conceptual research according to Rikhardsson & Yigitbasioglu (2018), the need to examine how IT systems in general are being implemented and their challenges has been included in order to be able to draw likenesses between the concepts and their challenges, since Big Data is closely related to the implementation of the tools and systems required to use Big Data analytics.

The tables below illustrate the technical and managerial challenges identified by different authors. The difference between the amount of identified challenges for the different categories, can be explained by the fact that even though some of the reasons of enterprise systems failures is due to technical challenges, these are not the main reason why IS fail. According to Davenport (1998) are the most significant challenges, managerial challenges, since organisations fail to reconcile the technological imperatives of the information systems (Davenport, 1998). To conclude, since these challenges are applicable on different cases of IT systems, is it therefore probable that these types of challenges also are valid on Big Data implementation.

References Challenges	ERPS	CRMS	Big Data
Poor data quality	Alshawi et al. (2004) Momoh et al. (2008) Momoh et al. (2010)		Ahmed et al (2017)
Sales employees pose poor technical skills		Ahearne et al. (2012)	
Data silos		Khamees (2013)	
Data security and confidentiality			Ahmed et al (2017)
Lack of database software			Ahmed et al (2017

Table 2: Technical challenges

References Challenges	ERPS	CRMS	Big Data
Inexperienced staff and limited training, or problem to find competent employees	Al-Mashari (2003) Cissna (1998)	Ramsey (2003) Nguyen et al. (2007)	McAfee et al. (2012) Ahmed et al. (2017)
Poor understanding of Business implication and requirements	Davenport (1998) Langenwalter (2000) Soh et al. (2000) Ehie & Madsen(2005) Momoh et al. (2008) Kogetsidis et al. (2008)	Kovacs (2006) Zimmer (2006) Nguyen et al. (2007)	McAfee et al. (2012) Morabito (2015)
Poor leadership	Cissna (1998) Al-Mashari (2003) Momoh et al. (2008)	Ramsey (2003) Nguyen et al. (2007)	McAfee et al. (2012) Ahmed et al. (2017)
Lack of top management support	Cissna (1998) Al-Mashari (2003) Ehie & Madsen (2005)	Ramsey (2003) Zimmer (2006) Nguyen (2007)	

Table 3: Managerial challenges

2.4   Research questions

From collecting the information in the theoretical framework, it can be understood that the implementation challenges can be divided into technical and managerial challenges. The following research questions has therefore been derived:

1. *What are the challenges to implement Big Data for decision-making purposes in large organisations?*
    1.1.   What are the technical challenges?
    1.2.   What are the managerial challenges?
    1.3.   What are the underlying reasons, that are generating the identified challenges?

3. Methodology

The Methodology section consists of the explanation of how the research has been conducted, including information about the case company, Volvo Cars Group, and how the interviews were conducted.

3.1   Research Strategy

The research strategy is the general plan on how to conduct the research in order to answer the research question(s) (Duignan, 2016). At the beginning of the research, after the topic of the study was determined, a case study approach was chosen in order to enable an understanding of the current situation of the usage of Big Data at a large organisation.

The case study approach comprises of the possibility to understand the nature of why and how the techniques, procedures and systems are used in practice. Case study can therefore be used in order to explore and illustrate the practices (Ryan, et al., 2002).

After the case study approach was chosen, contact was established with the case company (Volvo Car Group). The interviewees were contacted after an initial contact with Erik Severinsson (our contact at Volvo Car Group), who contributed with names of possible interviewees. Before the contact of interviewees, the review of the theoretical framework was conducted in order to gain an understanding of the fundamental concepts and the gap in previously made research. This enabled the authors to formulate relevant questions for the interviews. The research strategy developed for this study can be summarised by *figure 5*:
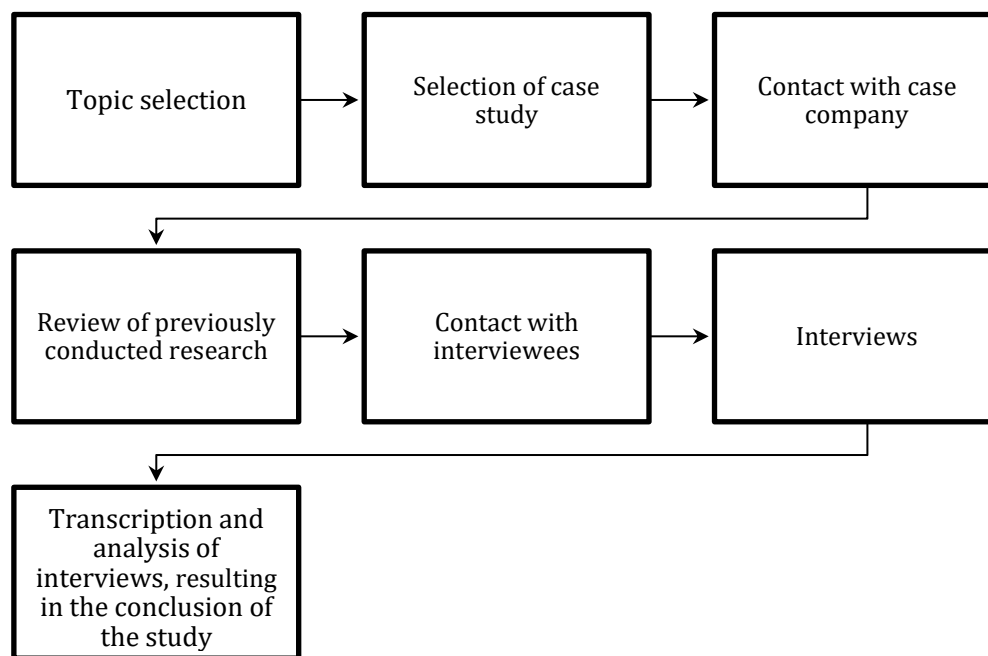

Figure 5: Research strategy

3.2   Research Design

The research design is how the research questions will be approached, including aspect such as how he data will be collected, how the interviewees were chosen (Duignan, 2016). For this study, an explanatory case study approach was selected, due to the aim to identify, as well as explain the causes behind the Big Data challenges. By conducting case study, an extensive understanding will be obtained to guide a profound analysis. In addition, the explanatory approach allows the authors to modify existing theories to be used in the study, since the aim is to generate theories to explain the current situation (Ryan, et al., 202). The need to modify existing theories have become evident from conducting the theoretical framework, where the lack of existing theories regarding the Big Data challenges has become clear.

The case company, Volvo Car Group (VCG), is used in order to examine their practices regarding Big Data, and thereby find explanations for the challenges, which could be applicable to other large organisations.

The research of this study has been conducted through investigating previously made research in order to gain a deeper understanding of the subject, and by conducting interviews at VCG. The study will be based upon a qualitative approach, due to a qualitative research focusing on the descriptive of the opinions, characteristics, behaviours etc. of the collected information (Duignan, 2016).

3.3   Data collection

3.3.1   Interviews

The empirical findings in this study are based upon information collected through several conducted interviews at VCG. The Interviews have been constructed as semi-structured interviews with open ended questions, allowing follow up questions when needed (Duignan, 2016). The interviews were recorded with the permission of the interviewee, and thereafter transcribed. This in order to minimise the risks of misunderstandings and omitting of important information.

A semi-structured approach means that the interviews has followed a pre-set questionnaire but has also allowed for follow up questions and some changes to the questions depending on the interviewee and his or her knowledge area (Duignan, 2016).

Examples of people to conduct interviews with, were firstly given to us by our supervisor at VCG, Erik Severinsson. The interviewees were thereafter contacted, and after the respective interview, we asked the interviewee if he or she had any suggestion of who to interview next. This approach allowed us to be able to conduct several interviews with employees with various positions at VCG, providing us with diverse knowledge and information.

The data from the interviews were analysed by firstly carefully transcribing the information, interview by interview. Thereafter were the information allocated into the different functions/departments to which it was connected. The information was thereafter compared to enable the identification of similarities and differences among the data.

3.4   Quality of the research

3.4.1   Validity

The concept of validity refers to the extent of which the information of the study reflects the reality of the situation. The validity concept can be divided into *internal* and *external* validity. The internal validity concerns the usage of internal controls for the research, while the external validity concerns the generalizability of the results of the study (Ryan, et al., 2002). To reduce the risk of a negative effect on the internal validity, the interview questions were prepared beforehand, in order to comprise of, for the study, relevant questions. In addition to this the recording of the interviews enabled the authors to rewind and confirm their recognition of the answers in order to be able to conduct an accurate description and analysis of the current situation. The external validity has been taken into consideration by the inclusion of several interviewees from different departments, in order to be able to base the analysis on more sources, and trying to maintain an objective approach to the data. Even though these measures have been done, due to the study being based upon a case study, there might be some restrictions to the possibility to generalise the results (Duignan, 2016).

3.4.2   Reliability

The reliability of the study refers to the findings, and their independence from the user. Meaning that the research results would be the same, if the research were to be conducted again, by someone else (Ryan, et al., 2002). In order to achieve a high reliability of the results in this research, eight interviews were conducted in order to gain information from several sources to be compared in its similarities and differences. The information from the interviews has thereafter been carefully analysed. Since the research is dependent on interviews and human knowledge, this might affect the reliability of the study, due to the interviewees changing their answers to some extent, if the study were to be conducted again.

3.5   Volvo Car Group and Big Data

Volvo Cars Group (VCG) is one of the world's leading car companies. The potential of using of Big Data, has become clear for VCG, even though they consider themselves as being in a starting point position, comparing to Google, Facebook and other 'Data Companies' (Samad, 2018). According Zerbino, et al (2017), Big Data and Big Data Analytics are transforming customer-facing industries. This information contributes to the fact that VCG need to create and have a

standardised designed system in order to control the usage of Big Data, and to be able to have a strategic manner of using Big Data in the future (Interviewee 2, 2018).

According to the annual report of 2016 for Volvo Car group, is their mission "*to be the world's most progressive and desired car company and to make people's lives less complicated*". The brand strategy of VCG is that everything that they do starts with people, they protect what's important for people and wants to make people feel special. Calling this "*Designed Around You*" and it is "*all about seeing, hearing and understanding people's needs and confirming through delivering experiences that exceeds their expectations and support them in their daily life*" (Annual report 2016, Volvo Car Group).

To be able to meet these goals in the future, is VCG focused on using Big Data, and is currently using Big Data in some specific projects. Big Data is regarded as an extremely important tool, especially since the car industry is rapidly changing, in aspects such as electrification, car sharing and autonomous driving. Furthermore, according to the interviewees, the CEO (Håkan Samuelsson) of VCG made an internal speech and illustrated that the organisation will deposit more of the budget into achieving better, and more aggregated systems, in order to enable easier access to the data. In addition to this, the CEO stated that the quality of data will also be emphasised in the future.

However, the work with Big Data has not been launched as a common way of working at VCG, thereby being at an early stage of using these tools. Since VCG have started to use Big Data and is in the early stage of using it, they have encounter several challenges which is interesting to investigate from a researching perspective (Interviewee 1, 2018).

4. Empirical Findings

The Empirical Findings consists of the identified challenges from the conducted interviews which aim to answer the research questions. The sub-chapters contain definitions, challenges, and opinions of using Big Data for decision making. The first four sub-chapters contains challenges within different functions while the last chapter contains challenges within four different projects. The information from the interviews will then be the foundation of the forthcoming sections.

4.1 Identified challenges

4.1.1 Global Finance Operations

The first interview conducted for this thesis was with a finance department director (*Interviewee 1*), considering himself as being data-driven and realising the value of data. Big Data could be used by examining different pools of data in order to find correlations, affecting the business outcome in order to make decisions affecting the organisation according to him. Combining different data sources is essential in order to understand the customers and to find correlations in order to improve employee satisfaction.

VCG is good at processing the internal data, e.g. cost tracking of cars across different systems, but the internal data has some challenges concerning the fact that the data has a silo structure, preventing the possibility to find correlating internal data. This can be exemplified by the reports concerning operational costs, having no data correlating to for example staff satisfaction, which might affect the operational costs. In contrast to the internal data process, the processing of external data (e.g. customer's credit history etc.), requires more improvements.

The finance department has, according to interviewee 1, the tools to implement Big Data, but not the time or the competent people to find the correlations, since the employees lack the knowledge and education in order to be able to use the Big Data tools.

Another important aspect in order to be able to implement Big Data, is to simplify the Big Data process, for example by being able to load the data without assistance from the IT department.

> *"When it comes to the timeliness of the data, the need for using the reposts into decision making are a lot faster than what the IT-people are able to provide us with the reports."*
> *(Interviewee 1)*

4.1.2   VCG Sales

According to *Interviewee 2*, the CFO at VCG Sales, does Big Data not automatically lead to better decisions for marketing, but it has the potential. Big Data can be considered as raw material and an essential element for the business:

> *"It is the insights derived from Big Data, the decisions and actions people take that makes all the difference, rather than data itself."*

A challenge to implement Big Data to improve the decisions at the sales department, is the large amounts of data available, but not the tools and competent sales employees who have strong technical skills in order make best use of the Big Data. The possibility to have quick access to data, as well as being able to handle the data and access the data analytics, are the most prominent improvements for the usage of Big Data. In addition to this, the timeliness of the data is of importance, since the requested report and data, consolidated and structured to make decisions must be current and up-to-date, and enabling support and contribution to the value chain, e.g. there is a need for a standardised way of processing Big Data.

Another challenge regards to poor data quality, the existing systems are quality based, containing all the data, with all its variance and product numbers. But when the data is to be reported into the group reporting system, the data must be aggregated. However, the aggregated data do not allow the users to examine the exact profits, expressed as market by market earnings, only market earnings on a certain level is possible to examine. There is therefore a need to be able to utilize Big Data in order to enable the reports to be broken down onto a detailed level, in order to enable correct judgements on a aggregated level, showing the variations in single parts on a lower level.

4.1.3   Global Customer Service

In the Global Customer Service department, one business controller (interview 3) and two directors (interview 4 and 5) was interviewed. *Interviewee 3* refers to the data coming from the customer surveys as Big Data, due to the large sample size originating from several thousand dealers. This in contrast with *Interviewee 5*, who regards Big Data as data from many various sources and in a larger volume than normal data (small data), stating that most employees at VCG have heard of the concept of Big Data.

Big Data can be highly used for decision making in marketing, for example for "*Targeting Marketing*" where Big Data currently is utilized to a high degree. Another current usage of Big Data is to identify the customer's preferences from what the customers "click on" on different websites.

24

"All the decision-making procedures can be included in the management control system, since it is where the collection and gathering of data is consolidated, and thereby make correct decisions." (Interviewee 4)

Big Data will be crucial for all businesses in the future in order to make correct decisions. VCG does not currently have the knowledge of how to use all the available data. Preventing VCG from being a data driven company, since the usage of Big Data varies between different departments, the Sales and Customer Service as being a department not using Big Data to a large extent.

Interviewee 3 states that there is a vast amount of data collected at the Global Customer Service, but there is a lack of knowledge of how to use it in decision processes. On the marketing side, a vast amount of diverse data is being collected, but only a fraction is used. Interviewee 5 believes that it is mixed regarding how managers trust the Big Data findings. If certain data do not exist, then previous knowledge and experience is of utmost importance. This is also the case when there is no time to collect the necessary data. In many cases, there are obstacles in order to access the required data. These obstacles can be summarised into three main challenges at the Global Service Department:

- **Challenge 1**: Big Data is too concentrated and aggregated, it is not possible to break it down into details,

- **Challenge 2**: Certain Data sources connections are missing, it would be preferable to be able to collect the different data sources into one Big Data warehouse and extract the necessary and required data for the specific tasks, and

- **Challenge 3**: Lack of someone that is responsible and who can help with data silos and data quality problems.

Regarding Challenge 1, when it comes to steering there is limited data regarding pre-gross margin in the data source. When the data is broken down in order to examine at the profit level, it is only possible to investigate this in different product groups.

To exemplify challenge 2, the data sources containing information about price data, sales data, and customer satisfaction data, could be gathered into one data source and thereafter extract the data in order to investigate whether there are any correlations between the data.

The last challenge is rather fundamental, referring to the need of having someone responsible to contact for problems regarding the data and its tools.

4.1.4 The Data Scientists Team

This subchapter contains interviews with three different data scientists (interview 6, 7 and 8). The data scientists work as internal consultants within R&D, product strategy, marketing etc.

Moreover, the project leader of each data scientists team is often considered as a senior data expert. The data scientists team are working close with the business professionals and help with making of decisions

According to *Interviewee 6* there is no simple definition of Big Data, for example, the volume of the data set and the capacity of the memory today is much larger than it was a few years ago. *Interviewee 7* argues, unlike interviewee 3 from Global Customer Service, that the information from customer survey is not Big Data. However, web data, diagnosis data and signal data generated by customers, are considered as Big Data. *Interviewee 8* explains further that there is no specific definition of Big Data within VCG, or even a common definition of Big Data at the different departments at VCG, the definition will be dependent on who you query.

Big Data is being used to predict selling volumes as well as which attributes are the most attractive for potential customers, enabling managers to identify the most popular car models. A large amount of external data is bought from suppliers (other brands). The usage of Big Data contributes with many opportunities, for example being able to discover new trends. In addition to this, by using Big Data it would be possible to know and understand the influences of the factors leading to the result, in order to increase finance control. According to interviewee 6, the larger a project is, the more data-driven it becomes, since a large project might be connected to higher risks and require more resources.

The access to external data is rather limited, with only some employees having access to the data. This is because, if everyone were to have a licence it would become excessively expensive. As for internal data, there are also some restrictions, for example regarding sensitive and confidential finance data from SAP.

According to interviewee 7, most employees at VCG are unaware of Big Data, resulting in the need of informing managers about the opportunities to use Big Data analyses in order to make decisions and changes. This since some managers are more inclined to use previous experience and intuition rather than data analysis. He concludes the interview with stating that:

"When some people in VCG say we are in the 'start point' of using Big Data, is because they don't know what we (as data scientist) are doing, there are many advanced projects within Big Data that we are doing right now, but it is not that everyone aware of it."

Interviewee 8 further identifies several challenges:

- **Inefficient data mining process**. For example, data from SAP is not directly accessible from the data warehouse, requiring additional data mining. It would therefore be preferable to enable extraction from the data warehouse directly.

● **Lack of communication between different Big Data project teams:** It would be preferable if different teams would communicate what they are working with, since it may be of interest to others, in addition to being able to use some else's knowledge.

● **Lack of a "simple conceptual system".** The challenge refers to the need of business and IT professionals being able to "speak the same language", in order to communicate better.

● **Lack of knowledge of how to use Big Data,** is the current main challenge, due to the lack of knowledge of how to integrate Big Data into the business. The problem is the need to integrate a data driven model for the business employees use to make decisions.

● **Lack of responsible person for data quality and data silos problems,** there is no one at VCG who is responsible for resolving the data quality problem. This even though there have been efforts made in order to raise the awareness of the value of data.

4.1.5 Cross functional projects

From the interviews, four different projects in which Big Data is being used for decision making have been identified: *product planning, predictive maintenance, residual value analysis,* and *customer behaviour analysis*. Each project generating benefits for VCG, but also encountering challenges to fully make use of those benefits.

Product planning is for example using predictive analysis for the US market. In the US, customers buy a 'speculate car' instead of a customized car. By analysing data from previous years, including data regarding the features of the car (colour, model, engine, seats, etc.) many conclusions can be drawn. The data scientist investigate which specifications makes the car sell faster. The outcome of the analysis is put into a system in order to automate and optimize the best features. Data restriction has been identified as the main challenge in this project. Especially financial data from SAP is restricted to some data scientists. It would therefore be preferred for the data scientists to have access to data from SAP, generating benefits to the prediction analysis for product planning. However, the other data scientist does not regard this as a challenge, since SAP is not restricted for him. Inefficient data mining processes is a big issue in this analysis when he needed data from SAP.

Predictive maintenance has a wide range of uses. By analysing correlations among historical data from workshops, it has been discovered that certain parts of the car have a higher risk of breaking under certain circumstances. By using predictive maintenance, the technician can ask the customer to come to the workshop before the part breaks. The predictive maintenance project reduces the warranty cost for VCG and at the same time, increases customer satisfaction. The high volume of available data affects the data quality negatively, since it might differ in its

collection. Another challenge mentioned by interviewee 8 is that VCG are not able to extract real-time data directly from the cars. Therefore, the project is restricted to using historical data.

Residual value analysis refers to predicting the future prices of the cars based on their value today. The information is integrated from various data sources, investigating which factors are affecting the prices in different markets. This type of analysis uses a large amount of external data from dealers all over the world. The collection of external data, the which data scientists have no control over, is regarded as a main challenge. From the perspective of the data scientists, this is highly important to take into consideration, since some data might be missing, and the external data is highly affecting the project.

VCG is also collecting data in regard to *customer behaviour*. For example, web-click data contributes with investigating what customers want through the configurations on their website in the car model configurator. It is important to observe customer behaviour through the configuration data especially in the US, since the customers from different states show vastly different interests of types of cars. However, in most of the cases, VCG still has limited usage of all the customer behaviour data in its database. Interviewee 2 argues the access to this data is restricted for most of employees, resulting in the organisation lacking in knowledge of how to use the data. In addition to this, interviewee 8 states that since the customer behaviour varies vastly, it is hard to measure, and the measurement systems is hard to set up, due to it being related to psychological facts, resulting in the performance of experiments in order to know what to measure.

4.2   Summary of main findings

4.2.1   Challenges in the cross functions

Functions	Type of challenges	Challenges
Global Finance Operations	Managerial	Inexperienced staff and limited training
	Technical	Data silos
		Inefficient Big Data processes
VCG Sales	Technical	Sales people pose poor technical skills
		Lack of Big Data tools
		Inefficient Big Data processes
		Poor data quality
Global Customer Service	Managerial	Misconception of Big Data definition
		Poor Business-IT communication
		Lack of knowledge of Big Data utilization
		Lack of trust of Big Data finding
	Technical	Data silos
		Data restriction
		Poor data quality
		Inefficient Big Data processes
The Data Scientists Team	Managerial	Lack of knowledge of Big Data utilization
		Poor partnership between different Big Data project teams
		Poor awareness of Big Data
		Poor Business-IT communication
	Technical	Data silos
		Data restriction
		Poor data quality

Table 4: Challenges in cross functions

4.2.2 Challenges in the projects

Projects	Type of challenges	Challenges
Product planning	Technical	Data restriction
		Inefficient Big Data process
Predictive maintenance	Technical	Poor data quality
Residual value analysis	Technical	Data restriction
Customer behaviour analysis	Technical	Data restriction
	Managerial	Poor awareness of Big Data
		Lack of knowledge of Big Data utilization

Table 5: Challenges in the projects

5. Analysis

In this section the information from the Empirical Findings, will be analysed in order to gain a deeper understanding and knowledge of the current Big Data implementation challenges at VCG in order to further illustrate and answer the research questions. The challenges are presented and divided into the two main categories: Technical, and Managerial challenges.

5.1   Big Data implementation challenges across functions

5.1.1   Technical challenges across functions

5.1.1.1   Data quality

Poor information quality can have a significant impact on an organisation and its operations (Strong et al., 1997). Three of the functions at VCG: sales, global customer service, and the data scientists team, have all identified poor data quality as one of the main challenges when working with Big Data. The different departments describe the challenge differently, but share the same result. Poor data quality creates difficulties and complications for the Big Data utilisation, this can be examined at the sales department. The sales department is experiencing problems with the sales data, due to the deficiency of being able to break down the information to smaller sub-categories or sub-groups, resulting in the employee's inability to gain a deeper understanding of the contents of the data. This can be connected to the lack of Big Data tools, being able to break down the aggregated data-sets. In addition to this, several interviewees state that there is no responsible employee, specialized in solving the data quality problem. The lack of a responsible person can be an explanation as to why the data quality is a current challenge, which must be resolved.

The data quality seems to be a reoccurring challenge at VCG, the data scientists argue that VCG collects substantial amounts of data, with the volume constantly increasing. Resulting in the extraction of relevant data becoming increasingly more difficult, due to the increased complexity to find and investigate which relevant data, this in addition to the requirement to focus on maintaining a high quality in data rather than solely including large amounts of data that might be irrelevant. Another reason behind poor data quality is due to the fact that the information is not being actively used to a large extent. The lack of active usage of the data, makes it difficult to notice the when and where the lack of quality occurs. If Big Data had been used to a larger extent, this problem would have been possible to discover at an earlier stage and thereby been solved faster, improving the quality.

5.1.1.2   Data silos

The storage of data in data silos, has been identified as a key challenge by four of the different functions. The Interviewee from the finance department explains that the challenge exists due to the diverse data sources (e.g. price, sales and customer satisfaction data) are not connected to a data warehouse. Causing challenges when attempting to find correlations between the different data. Finding these probable correlations is highly important, due to the high possibility of existence and the impact these correlations could have on the organisation and its operations, thus discover new business opportunities, as well as threats and weaknesses. The probable corrections would be possible to examine if the data sources were connected.

The authors brought up this challenge to the next interview, with the manager of the sales department. He regarded it as something to be investigated in the future, but not as something to be prioritised. The fact that this is not considered as a prioritised challenge, and the manager's lack of intent to resolve the challenge can be regarded as one explanation of why the challenge exists. In addition to this, an interviewee from the Global Customer Service stated that there has never been a responsible employee to contact regarding the challenge, which is a further explanation to why the data silos has not been solved.

According to the data scientists it would be preferable to, instead of having separate data sources, have a data warehouse, allowing the diverse data sources to be connected and thereby enabling the finding of important correlations between the data. This would thereby allow a better decision-making process, since the new correlations could impact the decisions and how the organisation operate. However, to do this requires a large amount of work, and the prerequisites of new databases and more advanced tools in order to ameliorate the data warehouse and link previously unconnected data sources.

5.1.1.3   Data restriction

Data restriction affects the organisation and is therefore another important identified challenge, as highlighted by the data scientists. The restrictions can be divided in two categories, restriction of *internal* and *external* data. The restriction to the internal data can be exemplified by the limited access to the financial information from SAP. The restricted access is creating problems in regard to the timeliness of the reports and complicates the Big Data process, preventing an effective Big Data implementation at VCG. Moreover, the external data is limited due to licenses, required in order to gain access, required to be bought by VCG. The expensive licenses contribute with increasing the barriers for VCG, since few employees will have access. Due to the restrictions, the extraction of restricted data, both internal and external, require the involvement of several different employees in order to gain access to the required data for conducting the analytics, causing the process to be high time consuming.

The data restriction challenge is highly relevant to the Data Scientists Team, probably due to the fact that they are the ones actually handling the data. Many of the other functions, do not use the data to the same extent by themselves, since they are ordering the conduction of the analyses from the data scientists.

5.1.1.4 The Big Data process

From conducting the interviews, it has become evident that there is a need for a more efficient process of handling and processing Big Data. The timeliness is an important factor, closely connected to the timeliness issue regarding to the data restriction, since the data restriction also affects the Big Data process. The problem with not being able to have access to the Big Data analyses, and thereby conduct the analysis is problematic. The analyses are time consuming due to the number of steps required, leading to an disinterest in using the analyses since the results/answers are not accessible quick enough. This due to the fact that decisions often have to be made quickly, and not after several weeks.

One of the main reasons for the time-consuming process seems to be the need of involvement of several different parties, due to the restriction of data access. Some of the data scientists cannot access to the data that is needed for creating models and performing analysis, and therefore have to ask other parties to give them the data and then conduct the analysis. As stated by interviewee 1, he would like to be able to conduct he analysis himself without the IT professionals, while the data scientists on the other hand believes there is a need for well-constructed teams, and more communication in order to create more efficiency.

The interviewees also argued for the need of having a responsible employee of helping and solving technical challenges, such as the data quality challenge, and connecting the data sources. This would in addition to investments in better aggregated systems make it easier to access the data, increase the data quality, and create a more efficient way of working with Big Data, and thereby streamline the Big Data process.

5.1.2 Managerial challenges across functions

5.1.2.1 Big Data definition

As mentioned by the literature, there are no general definition of Big Data, creating difficulties when employees define Big Data differently in the different departments within the same organisation. This is the case at VCG, since some employees states that they are working with Big Data, but other employees do not consider that specific data as Big Data. This becomes evident when comparing the information from the interviews conducted with the customer service department and the data scientists. Global Customer Service regards the data from customer survey as Big Data due to its volume, but this is contradicted by both the data scientists and the literatures, due to its lack of variety and velocity. Therefore, the challenge regarding the

definition of Big Data, be identified as a misconception of the circumscription or lack of knowledge of Big Data.

The lack of definition is confirmed by the data scientists, due to the rapid change in the development on how to handle the data. But each department still seems to have its own definition of Big Data, creating difficulties since the different definitions will contribute to misunderstandings.

5.1.2.2   Data-driven culture

Several interviewees from the different functions at VCG, consider the organisation to be at the *'starting point'* or in a *'early stage'* of using Big Data. The significant managerial challenge faced by VCG to implement Big Data is, that the organisational culture is not data-focused enough. Both the data scientists and the employees from the Global Customer Service and VCG Sales argue that the customer service and sales department have a limited data focused culture.

This can be seen by the fact that, according to one of the data scientist, there are few managers making decisions based on data-driven models, and in some cases, the attention paid to the Big Data findings are minor, and not used as a basis for decisions, instead choosing to trust previous experiences and intuitions. Even though a large amount of data has been collected at, for example, the sales department, the employees still lack the knowledge of utilize the information. This scenario explains to some extent why data-driven models are not being used to a larger extent. In addition, an explanation can be derived from the need of creating a 'Big Data' thinking among the employees. The difficulties connected with changing the way of thinking might be due to, for example, employees being reluctant to change from using IT tools they are familiar with. This is contradicted by the fact that another of the data scientist have a different opinion, and that VCG is not at the *'early stage'* of using Big Data, since there are many advanced active Big Data projects at VCG. One of the underlying causes for this contradiction, could be the lack of a clear definition of Big Data, and the lack of attention paid to these projects by the employees in general. To solve this issue, requires spreading the information in order to enable the employees to become knowledgeable about Big Data, and understand its organisational value. One of the main reasons for the lack of a data-driven culture at VCG could therefore be due to the fact that the employees generally are not knowledgeable concerning Big Data and its benefits.

5.1.2.3   Communication

From the interviews, it can be concluded that there appear to be a general lack of communication between departments and employees regarding Big Data. One important aspect correlated to the communication, is that the employees do not know who to contact regarding data concerns. This can also be confirmed by the fact that some interviewees stated the usage of Big Data being at a

starting point, and at some departments not used to a large extent. This is contradicted by the IT-department, stating that the cause for this statement, is due to some employees having no thoughts of what the data scientist are doing, or what the data scientists can contribute with.

The lack of communication can also be connected to, the lacking definition of Big Data. The lack of communication creates challenges such as misunderstandings, and frustration, due to employees using the same concept, but including different aspects, resulting in not discussing the same things, even though they believe they do.

Moreover, the need to improve the communication skills can also be identified from some of the interviewees stating the cause of the Big Data projects not being valued correctly, is due to some of the projects not being fully used or appreciated by the business professionals. The cause is likely due to some of the project leaders from the IT department, not being able to deliver the value of Big Data analysis appropriately.

The lack of communication between the different Big Data teams, not communicating their findings and knowledge from previous projects to the other teams, causes previously gained knowledge to not be used in other projects, as stated by the data scientists. A better communication between the teams, would therefore be a positive change, since it would enable the different teams to share knowledge and information. By allowing the diverse teams to know and gain access to the other team's knowledge would facilitate the usage of Big Data, since some information could be of interest to other departments, functions, employees etc.

5.2 Big Data implementation challenges of cross functional projects

5.2.1 Technical challenges in cross functional projects

5.2.1.1 Data quality

Poor data quality can be identified as a challenge to the predictive maintenance project. The cause of the deficient data quality is due to the diverse sources of data. This is due to the fact that every time a Volvo arrives to a workshop, the data is collected and integrated to the data warehouse after data mining work. Resulting in the database containing a vast amount of information and data, unnecessary for the predictive maintenance analysis, causing the extraction and handling of the required data to become more complex and time consuming. In addition to this, the lack of structure regarding how the data is collected, affects the data quality negatively.

Based on this, the data lacking in quality, will affect the models created by the data scientists. Even though the volume of data is large, which could to some extent compensate the deficient data, it could still result in the data scientist creating the wrong models to be used as basis of decision making. Furthermore, to use real-time data in this project can enhance the validity and reliability of the information, resulting in better decision making. However, the usage of real-

time data has technical restrictions, but if these technical obstacles could be overcome, it would highly affect how VCG could conduct their analyses of predictive maintenance.

5.2.1.2   Data restriction

Regarding data restriction, there are two distinct categories which are identified in several of the projects. The first category is concerning VCG's ability to acquire data from external sources, which is used for the residual value analysis. The second category is describing the internal data restriction to employees, which applies to product planning and customer behaviour analysis.

Regarding the residual value analysis, the difficulties in requiring all the needed external data, is due to the deficient standardisation of what data is collected and registered by the car dealer. In addition to this, the data considered important to the analysis changes over the years, even if there were to be a standardised data collection, some car dealers might not adopt to the new requirements, since the reaction time for these changes differs between regions and countries. This causes the absence of certain data required for the analysis, since some dealers do not collect the specific data.

The access to SAP is restricted, causing some problems, since the data might be useful for the prediction analysis for product planning. By connecting the financial data to the sales data could facilitate analyses regarding what cars to produce and sell. One explanation for this phenomenon is the lack of someone responsible to handle the task of connecting certain data into a data warehouse. Having the data centralised would provide all interested employees to actively use the data in order to conduct useful analyses and finding new correlations. This is, however, not the current situation at VCG, resulting in the accessing and processing customer behaviour data as limited.

5.2.1.3   The Big Data process

The Big Data process is inefficient, mainly due to the data mining being high time-consuming, and has therefore been identified as a key technical challenge, for example for the prediction analysis project. The main cause is due to some of the data not being directly accessible from a data warehouse. Resulting in the requirement of conducting additional data mining by a data engineer, before the data can be used for analytical purposes. Additionally, in the data analysis teams, the responsibilities of the employees are not clearly defined, resulting in not knowing who is responsible for each task, and who to contact in order to gain access and/or solve diverse problems. For example, should the data engineers focus solely on extracting information from the data sources while the data scientists should only on be concentrating on the modelling of the data, and create the model to be used for the decision making? To conclude, if the distribution of duties and accountabilities are not created within the team, can therefore result in an inefficient Big Data process.

5.2.2   Managerial challenges in cross functional projects

There has been large amount of data collected for the customer behaviour analysis, however the employees are short of knowledge about how to utilize the gathered information. The knowledge of how to use the data is highly important to the organisation, since it will affect the decisions being made, and thereby the operations of the organisation. The lackadaisical approach of understanding the data and its value, could be in likeness to the findings in regarding the managerial challenges across functions, be due to a lack of communication from the managers to the employees. A good communication is highly important in order to convey the reason why something is valuable and necessary. Another plausible explanation for this scenario is the lack of combining business employees and specialised data scientists, in order to enabling a discussion regarding the data. A discussion and examination of the data in a business case with, for example, customer behaviour data, could contribute to a deeper knowledge and understanding for the potential benefits and value the data could contribute with to the organisation

Even though the managerial challenges in cross functional projects have not been highlighted in the interviews, it can be concluded that it is likely that the challenges identified for across functions is applicable on the cross functional projects as well. The lack of definition of Big Data, is probably a contributing factor to the lack of understanding the potential value of using the data and how it can affect the organisation. This is also closely connected to the data focused culture, if the culture is not data-driven, the likeliness of using Big Data for analyses will decrease.

5.3   Key findings of the analysis

To summarise, the cause behind highlighting these key challenges is due to these issues prevent the usage and to effectively implement Big Data in order to support the management and the organisation in the decision process. These challenges affect the continuation of the development of the Big Data tool, as well as, if not addressed, decreases the quality of the decisions made while the Big Data tool is not fully developed.

The key findings of the *technical challenges* can be summarised and identified as: *poor data quality, data silos, data restriction,* and *an inefficient Big Data process*. These challenges are critical to overcome due to their effect on the quality of the outcome of the Big Data analysis in form of decision support. If not addressed properly, these challenges risk the decisions process resulting in incorrect and/or delayed decisions (due to time-consuming process to perform Big Data analysis).

The *managerial challenge's* key findings are identified as: *the definition of Big Data,* due to the misconception of what Big Data is, a less than ideal *data driven culture,* as well as *communication issues* between business and IT personnel, as well as the lacking communication

between the different IT projects. The most critical challenge is the communication challenge, since the effects of this issue, affects all the other challenges as well. The importance of good communication can be illustrated by the Big Data process being underdeveloped due to deficient communication between the IT projects. Another example derived from the communication challenge, is the conclusions and decisions made with Big Data analysis as support, will be negatively affected due to the misalignment between the business and IT functions.

The second critical managerial challenge is to create a more data-driven culture. Some of the interviewees has identified that the culture at VCG to be data driven, while others have identified it as less data driven, dependable upon the functions. The cause of the importance of the culture, is due the culture infuses the entire organisation, and with a data-driven culture more efforts and value will be incorporated when conducting Big Data analysis as decision support tool, thereby increasing the quality of the decisions being.

6. Discussion

The discussion section will highlight the findings from the Analysis section, together with information from the Theoretical Framework in order to find similarities and differences among the findings and the theories used. The discussion will make use of the two tables under for each subchapter. In the summarised table, challenges of ERPS and CRMS from Theoretical Framework as well as Big Data challenges coming from Empirical Finding will be presented.

6.1 Technical challenges concerning the decision process

In the *Table 6* below, the identified technical challenges both from the theoretical framework as well as from the empirical findings are presented. Half of the findings are shared between Big Data and ERPS/CRMS while the other half is unique to the empirical finding. This is because, as highlighted in theoretical framework, in different articles, the authors have less emphasis on technical issues for ERPS/CRMS implementation processes (Huang et al, 2003; 2004). Due to these IT tools have been implemented and used for a long time, the technical aspects of the tools are more mature.

First of the challenges is connected to poor data quality and is shared between Big Data and ERPS. The ERPS poor data quality challenge is connected to data accuracy, meaning that the data is inaccurate or not immediately accessible (Alshawi et al. 2004). Youngberg et al (2009) highlights that one of the ERPS challenges are *"getting the right information to the right person at the right time and in a usable form"*. The poor data quality is a highly important aspect that must be resolved in order to enable a correct basis for steering and decision making in an organisation. Since if the decision were to be made based upon data with inadequate quality, this might lead to incorrect decisions. This challenge is also prevalent from the interviews, where the data according to some interviewees, is difficult to handle, as well as, occasionally, not being in a usable form, for example, in the customer behaviour analysis project. The Big Data is not always immediately accessible, hence the need for the data scientists to contact the data engineers, in similarity to the information from the theoretical framework. The empirical findings however add an additional important aspect of data quality relating to the problem of some financial and sales data being too aggregated, and not possible to be broken down to a lower level of aggregation while maintaining sufficient details in order to perform the sought after analyses. In order to be able to use Big Data for management decisions, the data must be possible to break down, since it will enable understandability and knowledge of why and how certain decisions must be made.

The second challenge shared between CRMS and Big Data is "sales employees with poor technical skills". The lower technical expertise possessed by the sales employees, affects the quality of the available information for managers to base their decisions upon, since it affects

how the collected data is collected, incorporated, and extracted. This is also connected to the finding of the employees not being fully aware of Big Data and its value to the organisation. In addition to this the lack of knowledge and understanding of how Big Data can improve the steering and decision process at the organisation. By improving the knowledge and understanding of the need for the data, and the required technical skills to be able to use the analytical tools, will facilitate the decision making, since employees will be able to use that available tools.

Moreover, the challenge regarding data silos in CRMS, as stated by Khamees (2013), is likely very similar to the data silos, faced by VCG. The data collected in CRMS is difficult to share with different data sources or applications (making it a data silo). This can be discovered from the empirical findings, where the need of being able to share data collected from diverse sources, as well as being able to connect the diverse information. The data silos are preventing the new opportunities of Big Data which could be discovered through finding new correlations between the different data. Moreover, by finding new correlations, the operation of the organisation could be improved due to the possibility of making more correct decisions based on the data.

The inefficient data processing challenge regarding Big Data, is connected to the fact that Big Data handles unstructured and semi-structured data in a larger extent than ERPS and CRMS. Those types of data forms require far more data mining work in order to make the data useful for analysis and drawing conclusions. The complexity of the works makes the processes of Big Data analytics highly time-consuming. The amount of time required to conduct the analytics, affect the management and their willingness to use Big Data analytics for the decision process, since time is of importance when making decisions. It is thereby highly important, in likeness with the literature, to have a standardised and efficient Big Data process.

The problem of the *"lack of tools"* is also applicable to only Big Data. This can be explained by the Big Data tool not being developed and used for as long period as ERPS and CRMS. Indicating that Big Data requires more resources in order to be useful as a tool for organisations. This challenge should be reduced as the continuation of the development of the Big Data technology. The lack of correct tools in order to conduct Big Data analytics affects the possibility to use Big Data for steering and decisions, since if the organisation do not have the tools, it does not matter whether they want to use Big Data and understand its value, since the organisation do not have the possibility to use it without the tools.

The challenge of data restriction applying to Big Data, affects the decision process. From the empirical findings it can be understood that the most prominent problem, concerns the access to external data. However, the external data is difficult to control by the IT department, it is expensive and sometimes might be inaccessible as discovered in the empirical finding, due to some data not being collected and/or due to privacy reasons. This differs from CRMS and ERPS, since these systems mainly handles and processes internal data.

Challenges	ERPS	CRMS	Big Data
Poor data quality	✕		✕
Sales people pose poor technical skills		✕	✕
Data silos		✕	✕
Inefficient Data processing			✕
Lack of tools			✕
Data restriction			✕

Table 6: The identified technical challenges

6.2 Managerial challenges concerning the decision process

In *Table 7* below, the identified managerial challenges, faced by Big Data, CRMS, and ERPS are presented. As can be seen, not all challenges mentioned in the theoretical framework are faced by all IT tools, this is also confirmed by the empirical findings. One main reason behind the significant differences between ERPS and CRMS compared to the Big Data findings can be due to the fact that the information in the theoretical framework is combined from several diverse studies, but the Big Data findings is based only upon the findings in this study. But as more and more studies will focus on the Big Data implementation challenges, in the future, the gap between the different tools will most likely decrease. The other reason may be due of the existing Big Data studies having more emphasis on the technical aspects rather than the managerial challenges within the research field of Big Data implementation challenges.

One challenge shared between all IT-tools is, the limited training and inexperience among the employees of using the tool. Huang et al., (2003) argues that ERPS is not easy to use even for highly educated employees, with excellent IT skills. This is also a fact in the CRMS implementation (Nguyen et al., 2007). This situation can be examined in the empirical findings, since many employees are excluded from the Big Data procedures, besides the IT department, resulting in the employees not being well trained in order to use Big Data, which the data scientists states is the main challenge in order to acquire a data-driven culture. However, this is not unusual, since when a new IT-system or tool is implemented, there will always be employees that is not used to working with the new tool, as well as having limited training in the beginning. The finite training of the employees in using Big Data, highly affects the possibility of using Big Data in the decision process, as well as in how to conduct the financial steering. Since if the

employees and managers do not have the skills to use the tools, they cannot, in likeness with not having the correct tools, use Big Data in the decision process at the different functions.

In the empirical findings, a challenge connected to poor communication has been identified, which has not been identified in the theoretical framework. One possible reason for this is that ERPS and CRMS users (business professionals) are less dependent on the help from IT professionals. In the Big Data process, the business employees rely heavily on IT employees due to the complicated models needed to be created, as well as the analysis being complicated to conduct. In the ERPS and CRMS, the core functions of the tools can be handled by the trained business employees themselves (Momoh et al., 2010; Ahearne et al., 2012). Therefore, more communication is needed between IT and business professionals in order to use Big Data, in order to create a better basis for decisions and steering. This challenge can be reduced in the future, by simplifying the way to use Big Data, by having a better Big Data process.

The understanding of the business implications and requirements of both ERPS and CRMS, is highlighted as a major challenge, due to the fact that businesses might not fully understand, how the information systems affect the business as a whole. One explanation for this, as stressed in the theoretical framework, is that organisations are trying to implement these systems in a rapid speed without understanding the tools (Momoh et al., 2010). This was not noted as a challenge in this thesis, in regard to Big Data, and might be due to the Big Data tools being more connected to the decision-making process, and more defined than the processes connected to ERPS and CRMS. Another explanation to this could be derived from the empirical findings, since, according to the interviewees, the business professionals set up the main tasks and requirements, and inform the data scientists of what they want to achieve, and thereafter discuss the requirements and how the data scientists could assist with Big Data as a support tool for the decision process. Therefore, will the addition and implementation of Big Data not affect the businesses employees' way of working in the same extent as implementing ERPS and CRMS, and will instead be used as an additional source of information.

The challenges of poor leadership as well as lack of executive support, as highlighted by the theoretical framework, is evident when implementing ERPS and CRMS. The need of executive support is highlighted by Howson (2013), as a fundamental aspect in order to succeed with any BI-program at the organisation. Without support and leadership when implementing Big Data in an organisation, the implementation will not be well integrated in the organisation, thereby affecting its effect for decision making negatively. However, this is not the case in VCG. The leader of each Big Data project is often considered as senior data expert who collaborates closely with business experts and work towards the aligned goal. Furthermore, all the interviewees strongly believe that the executive team in VCG gives an increasing attention on Big Data and sufficient budgets to various IT projects. The CEO of VCG illustrated that the organization plans to set up a higher budget into achieving more aggregated and advanced systems in order to be easier to access different kinds of data. Therefore, the lack of leadership and executive support have not been identified as a case with Big Data, and this is most probable connected to the fact

that VCG is very enthusiastic and is encouraging the use of Big Data, hence giving the support and leadership in order for the implementation to be managed in a good way. This in order to be able to use Big Data to better support the steering and different decisions conducted at VCG.

In contrast is the challenge of "poor awareness of the tool" only prevalent in the context of Big Data. The theoretical framework does not cover this challenge, but it becomes evident from the interviews, that there is a lack of knowledge about how Big Data is used. The lack of knowledge of how to use Big Data affects its usage for decisions negatively, since if the users do not know how to utilize the information, they will use it to a lesser degree. It is therefore highly important that the employees receive training about the value and how to use Big Data in order to improve the organisation and the decisions made by managers. The lack of knowledge of how to use Big Data might be due to the concept of Big Data being a fairly new concept, and it currently being promoted heavily by the management in many organisations. This means that many employees know the word Big Data but do not have a clear vision of what Big Data really means, and how it can support the organisation.

Challenges	ERPS	CRMS	Big Data
Inexperienced staff and limited training	✕	✕	✕
Poor Business-IT communication			✕
Poor understanding of business implication and requirement	✕	✕	
Poor leadership	✕	✕	
Lack of top management support	✕	✕	
Poor awareness of the tool			✕

Table 7: The identified managerial challenges

7. Conclusion

The concluding section of this thesis, the Conclusion, will summarise and explain the main findings, and contributions, regarding the challenges faced when implementing Big Data, derived from this paper. The section will be concluded with topics for further research within this field, and recommendations to be used in practice.

From conducting this research, it has become evident that, there are some key challenges faced by organisations when implementing Big Data. The main findings from this study are that there seems to be more technical rather than managerial challenges when implementing Big Data. This differs from the implementation challenges faced by ERPS and CRMS, where there seems to be more managerial challenges to consider then technical.

This thesis contributes to the existing research, by being based upon empirical findings, rather than conceptual research, which have been the foundation for previously conducted research (Rikhardsson & Yigitbasioglu, 2018). By using a case company for our research, we have been able to examine the challenges faced by large organisations in practice, and thereby gain a deeper understanding of the current situation. Another contribution of the thesis is the identification of gaps between Big Data implementation challenges compared to other IT tools as well as finding gaps within existing Big Data research. From conducting this research, the diverse implementation challenges between ERPS, CRMS, and Big Data, have been highlighted, and shown that it is not possible to directly transfer all IT tools implementation challenges to Big Data, even though Big Data is an IT tool. Some of the challenges faced by CRMS and ERPS are not perceived as a challenge for Big Data and vice versa. This can be derived based upon the information from the theoretical framework. The study confirms some of the challenges of other IT tools (ERPS and CRMS), to be applicable to Big Data, but it is also noted that even if the systems share some of the challenges, most of the Big Data implementation challenges were unique to Big Data. This was surprising, since the challenges shared by ERPS and CRMS, were rather equally shared, and would therefore be likely to be shared with Big Data as well. It was also surprising due to the fact that the theoretical framework was based on several studies as well as one meta study about this topic and still some challenges for Big Data had not been found in the ERPS or CRMS implementation.

From conducting this research, it has become evident that the implementation of using Big Data for decision processes for large organisations, contains several challenges that must be addressed in order to enable a successful and efficient usage of Big Data. The challenges differ in character and underlying factors for their appearance, but the most important challenge to address is the need of good communication. If the communication is well developed among all affected parties, such as IT employees, management, and business employees, i.e. all potential users of Big Data, the remaining challenges can be solved through communication.

The conclude this thesis, the answers to the research questions presented in the theoretical framework, could be summarised by the following: The technical challenges, summarised in Table 6, shows the most prevalent challenges to be poor data quality, data silos, inefficient Big Data process, as well as data restriction. These challenges do all contribute to a lower level of utilization of Big Data as decision support tool. The managerial challenges in Table 7, indicate the most prevalent challenges to be inexperienced staff, limited training and poor communication/teamwork between the business and IT departments. The reasons behind the occurrence of these challenges are not easy to summarise, due to the diverse underlying causes (as discussed in chapter 5). However, the two most prevalent reasons for these challenges are due to Big Data being a fairly new tool which has not been fully developed, resulting in foremost the technical challenges, but also causing the higher need for a collaboration between the IT and business departments due to the complexity of performing the Big Data analysis and aligning the analysis with the business goal.

7.1   Recommendation for practice

To conclude this thesis, some recommendations for practice will be provided, based on the theory and empirical findings, in order to help large organisations to overcome Big Data implementation challenges. The recommendations are sectioned into two parts, the first part focusing on solving the technical challenges of the Big Data process (hence improving the efficiency), while the second part of the recommendation aims to improve the effectiveness of the tool.

The first recommendation is to create either a department, or a work group responsible for Big Data and its implementation challenges, consisting of both IT-professionals, and business-professionals in order to optimise the process from the two different perspectives. This recommendation is motivated, by the finding from the theoretical framework that by allowing different employees to partake their different needs could be taken into account. In addition to this the generating of a department or work group, allows the employees from diverse departments to to know who to contact regarding problems or suggestions on how to make the work more efficient. Secondly, this will also improve the teamwork between the IT and business departments in order to overcome the challenges in a more effective way, this is also closely connected to the possibility to consider the different needs of the diverse departments. These challenges were clearly highlighted in the empirical findings and can be solved through this recommendation for practise.

The second recommendation regards the prioritising of what challenges to have primary focus on. Based on the empirical findings, the demand of Big Data analysis can be regarded as fairly high, but the time consumption to produce the analyses, and/or low trustworthiness of the data, might be a cause for the low usage of Big Data. These problems could result in a lack of interest of the Big Data analysis. This causes the development of the Big Data process to slow down. To

overcome this challenge, it would be recommended to focus on the challenges currently reducing the demand for Big Data as a decision support. These challenges can be identified as:

1. Creating a data-focused culture to increase the demand for Big Data analysis.
2. Develop a more efficient Big Data process as well as improving the data processing in order to create timelier results.
3. Improve the business and IT department collaboration (preferably through recommendation 1) to make the Big Data analysis result more applicable as decision support.

7.2   Future Research

The research regarding Big Data and its potential value contribution to organisations, are still very limited, especially regarding empirical research (Rikhardsson & Yigitbasioglu, 2018). From conducting this research, the study has mainly focused on the challenges concerning the implementation of using Big Data in large organisations, and the potential positive contributions of Big Data. To gain a deeper understanding of the contributions of Big Data, it would also be interesting to examine the potential negative aspects of using Big Data. Since it is likely that there could be negative aspects to consider in addition to the positive aspects. In addition to this, another interesting area of research would be to examine what extent the implementation of Big Data actually contribute to a better performance of the organisation. This in order to understand the impact of Big Data in practice, as well as if the size of the organisation and the industry affects the implementation of Big Data, since this study only focused on large organisations, due to our case company.

In addition to these research areas, it would be interesting to examine, how Big Data is being used in organisations after being fully implemented. For example, how is it used when analysing revenue, sales, and/or forecasting? To further develop this area, research regarding how the increased access to information affects the performance of the management accounting tasks, would be of interest.

References

Ahmed, V., Tezel, A., Aziz, Z. and Sibley, M., 2017. The future of big data in facilities management: opportunities and challenges. *Facilities*, *35*(13/14), pp.725-745.

Ahearne, M., Rapp, A., Mariadoss, B.J. and Ganesan, S., 2012. Challenges of CRM implementation in business-to-business markets: A contingency perspective. *Journal of Personal Selling & Sales Management*, *32*(1), pp.117-129.

Al-Mashari, M., 2003. Enterprise resource planning (ERP) systems: a research agenda. *Industrial Management & Data Systems*, *103*(1), pp.22-27.

Alshawi, S., Themistocleous, M. and Almadani, R., 2004. Integrating diverse ERP systems: a case study. *Journal of Enterprise Information Management*, *17*(6), pp.454-462.

Volvo Car Group. *Annual report 2016.* [Online] Available at: https://assets.volvocars.com/intl/~/media/international/annualreports/2016/static/pdf/annualreport2016.pdf [Accessed: 10 Apr. 2018]

Anon, 2006. Oracle Refreshes PeopleSoft CRM Software: CRM 9 offers usability enhancements and tighter links to Oracle's Fusion middleware. (customer relationship management)(Brief article). *InformationWeek*, pp. InformationWeek, August 14, 2006.

Anon, 2017. Big Data in the Danish industry: application and value creation. *Business Process Management Journal*, 23(3), pp.645–670.

Boyd, D. & Crawford, K., 2012. CRITICAL QUESTIONS FOR Big Data. *Information, Communication & Society*, 15(5), pp.662–679.

Butterfield, A., Ngondi, G. E., & Kerr, A, 2016. *A dictionary of Computer Science*. Oxford University Press.

Chen, S.G. and Lin, Y.K., 2009. On performance evaluation of ERP systems with fuzzy mathematics. *Expert Systems with Applications*, *36*(3), pp.6362-6367.

Cissna, T., 1998. ERP software implementation brings pain with its gains. *Electric Light & Power*, *76*(11), pp.43-44.

Daintith, J. & Wright, E., 2008. *A Dictionary of Computing*, Oxford University Press.

Data & Analytics Research, 2016. *International Data Group.* [Online] Available at: https://www.idg.com/tools-for-marketers/tech-2016-data-analytics-research/ [Accessed 12 Apr. 2018].

Davenport, T.H., 1998. Putting the enterprise into the enterprise system. *Harvard business review*, *76*(4).

Davenport, T.H. and Dyché, J., 2013. Big data in big companies. *International Institute for Analytics*, *3*.

Davenport, T.H, 2014. Big data at work: dispelling the myths, uncovering the opportunities. *harvard Business review Press.*

Davenport, T.H, 2014. How strategists use "big data" to support internal business decisions, discovery and production. *Strategy & Leadership*, *42*(4), pp.45-50.

Dechow, N. and Mouritsen, J., 2005. Enterprise resource planning systems, management control and the quest for integration. *Accounting, organizations and society*, *30*(7-8), pp.691-733.

De Mauro, A., Greco, M. and Grimaldi, M., 2016. A formal definition of Big Data based on its essential features. *Library Review*, *65*(3), pp.122-135.

Doyle, Randy, 2013. Avoid the IT sinkhole: new-system implementation has challenges, essentials for success.(Business Matters). *LP/Gas*, 73(10), p.45.

Duignan, J., 2016. research strategy. *A Dictionary of Business Research Methods*, pp.A Dictionary of Business Research Methods.

Ehie, I.C. and Madsen, M., 2005. Identifying critical issues in enterprise resource planning (ERP) implementation. *Computers in industry*, *56*(6), pp.545-557.

Frizzo-Barker, J., Chow-White, P.A., Mozafari, M. and Ha, D., 2016. An empirical study of the rise of big data in business scholarship. *International Journal of Information Management*, *36*(3), pp.403-413.

Gandomi, A. and Haider, M., 2015. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), pp.137-144.

George, G., Haas, M.R. and Pentland, A., 2014. Big data and management. *Academy of management Journal*, *57*(2), pp.321-326.

Global data management benchmark report, 2017. *Experian.* [Online] Available at: http://www.experian.com.my/wp-content/uploads/2017/12/2017-global-data-management-benchmark-report.pdf [Accessed 20 Apr. 2018].

Granlund, M., 2011. Extending AIS research to management accounting and control issues: A research note. *International Journal of Accounting Information Systems*, *12*(1), pp.3-19.

Hendricks, K.B., Singhal, V.R. and Stratman, J.K., 2007. The impact of enterprise systems on corporate performance: A study of ERP, SCM, and CRM system implementations. *Journal of operations management*, *25*(1), pp.65-82.

Howson, C., 2013. *Successful business intelligence: Unlock the value of BI & big data*. McGraw-Hill Education Group.

Huang, A., Yen, D.C., Chou, D.C. and Xu, Y., 2003. Corporate applications integration: challenges, opportunities, and implementation strategies. *Journal of Business and Management*, *9*(2), p.137.

Huang, M.H., Wang, J.C., Yu, S. and Chiu, C.C., 2004. Value-added ERP information into information goods: an economic analysis. *Industrial Management & Data Systems*, *104*(8), pp.689-697.

Interviewee 1, Volvo Car Group. 2018-2-14

Interviewee 2, Volvo Car Group. 2018-2-19

Interviewee 3, Volvo Car Group. 2018-2-22

Interviewee 4, Volvo Car Group. 2018-2-27

Interviewee 5, Volvo Car Group. 2018-3-1

Interviewee 6, Volvo Car Group. 2018-3-8

Interviewee 7, Volvo Car Group. 2018-3-14

Interviewee 8, Volvo Car Group. 2018-4-20

Khamees, R.A., 2013. Overcoming the Challenges of Implementing LMS, ERP and CRM in Traditional Business Environments: Survey. *IRACST-International Journal of Research in Management & Technology (IJRMT), ISSN*, pp.2249-9563.

Kogetsidis, H., Kokkinaki, A. and Soteriou, V., 2008. Enterprise resource planning implementation in the retail sector: a case study on the effects of the implementation of an ERP

system in Cyprus. *Journal for International Business and Entrepreneurship Development*, *3*(3-4), pp.270-283.

Kovacs, J., 2006. How to embrace e-business and make it succeed, Destination CRM, Viewpoint.

Labrinidis, A. and Jagadish, H.V., 2012. Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, *5*(12), pp.2032-2033.

Langenwalter, G.A., 1999. *Enterprise resources planning and beyond: integrating your entire organization*. CRC Press.

Mabert, Soni & Venkataramanan, 2003. Enterprise resource planning: Managing the implementation process. *European Journal of Operational Research*, 146(2), pp.302–314.

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

Marx, V., 2013. Biology: The big challenges of big data.

McAfee, A., Brynjolfsson, E., Davenport, T.H., Patil, D.J. and Barton, D., 2012. Big data: the management revolution. *Harvard business review*, *90*(10), pp.60-68.

Momoh, A., Roy, R. and Shehab, E., 2008. A work breakdown structure for implementing and costing an ERP project. *Communications of the IBIMA*, *6*(15), pp.94-103.

Momoh, A., Roy, R. and Shehab, E., 2010. Challenges in enterprise resource planning implementation: state-of-the-art. *Business Process Management Journal*, *16*(4), pp.537-565.

Morabito, V., 2015. *Big Data and Analytics: Strategic and Organizational Impacts Springer International Publishing Switzerland*.

Nguyen, T.H., Sherif, J.S. and Newby, M., 2007. Strategies for successful CRM implementation. *Information Management & Computer Security*, *15*(2), pp.102-115..

Parmar, R., Cohn D.L. and Marshall, A., 2014. Driving Innovation through Data. *IBM Institute for Business Value*.

Quattrone, P. and Hopper, T., 2005. A 'time–space odyssey': management control systems in two multinational organisations. *Accounting, Organizations and Society*, *30*(7-8), pp.735-764.

Ramsey, S., 2003. Introduction: strategy first, then CRM. *The Ultimate CRM Handbook, McGraw-Hill, New York, NY*, pp.13-17.

Rikhardsson & Yigitbasioglu, 2018. Business intelligence & analytics in management accounting research: Status and future focus. *International Journal of Accounting Information Systems*, 29, pp.37–58.

Rossi, R. and Hirama, K., 2015. Characterizing big data management. *Issues in Informing Science and Information Technology*, *12*(unknown), pp.165-180.

Russom, P., 2011. Big data analytics. *TDWI best practices report, fourth quarter*, *19*(4), pp.1-34.

Russom, P., 2013. Managing big data. *TDWI Best Practices Report, TDWI Research*, pp.1-40.

Ryan, B., Scapens, R.W. & Theobald, M., 2002. *Research method and methodology in finance and accounting* 2. ed., London: Thomson.

Soh, C., Kien, S.S. and Tay-Yap, J., 2000. Enterprise resource planning: cultural fits and misfits: is ERP a universal solution?. *Communications of the ACM*, *43*(4), pp.47-51.

Swift, R.S., 2001. Accelerating customer relationships: Using CRM and relationship technologies. *Prentice Hall Professional.*

Tesfaye, B., 2017. What is the influence or Big Data and Analytics on Management Control System.

Themistocleous, M., Irani, Z., O'Keefe, R.M. and Paul, R., 2001, January. ERP problems and application integration issues: An empirical survey. In *System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on* (pp. 10-pp). IEEE.

Vloet, L., 2016. Influence of big data and analytics on management control. Why changes in management control by means of big data and analytics are not achieved yet.

Ward, J.S. and Barker, A., 2013. Undefined by data: a survey of big data definitions. *arXiv preprint arXiv:1309.5821*.

Warren Jr, J.D., Moffitt, K.C. and Byrnes, P., 2015. How Big Data will change accounting. *Accounting Horizons*, *29*(2), pp.397-407.

Winter, S. 2011. Implementation. 1157-1170.

Youngberg, E., Olsen, D. and Hauser, K., 2009. Determinants of professionally autonomous end user acceptance in an enterprise resource planning system environment. *International journal of information management*, *29*(2), pp.138-144.

Zerbino, P., Aloini, D., Dulmin, R. and Mininno, V., 2018. Big Data-enabled Customer Relationship Management: A holistic approach. *Information Processing & Management*.

Zimmer, J., 2006. Be ready to take the heat. *Destination CRM. com, Viewpoint, available at: www. destinationcrm. com*, *15*, p.2007.

Appendix - Interview questions

The following interview questions were used as a guideline to collect information from the interviewees.

Big Data understanding
1. How would you define Big Data?
2. What kind of Big Data do you use in your work and where does the data comes from?
3. How does VCG engage in Big Data, from your perspective?
4. Are Big Data technologies incorporated in all business processes of the organization?
5. What is the organizational strategy for managing of Big Data in the future?
6. What are the success criteria for using Big Data?
7. Have you seen changes in the use of Big Data in the past few years in VCG, and if so what has changed?

Benefits of Big Data
1. What are the benefits of using Big Data in your daily activities?
2. Does Big Data play a role in decision making?
3. Does the use of Big Data helps you to discover new things?
4. How do you see the company's future in relation to usage of Big Data?

Challenges of Big Data
Technical
1. Is there a sufficient infrastructure to make a good use of Big Data in VCG? (analysing tools, storage capacity, structure etc.)
2. Can Big Data be collected on real-time manner and available to users? If yes, to what extent do you rely on this type of data for decision making purposes?
3. How does data warehouse looks like in VCG ? Any challenges to connect different data sources?

Managerial
1. Does the top management support the use of Big Data? What is the attitude of the top management regarding Big Data?
2. Do managers trust data findings in general, from your perspective?
3. Is there a cultural, structural shift in VCG because of Big Data?
4. Is using Big Data promoted/encouraged across all organization and by all employees? If yes, how and to what extent?

Qualification of employees

1. Do you think employees in VCG are aware of Big Data and know how to use it?
2. Do you get extra training in VCG in order to use Big Data?
3. Do all employees have access to all types of data? How is access to various data granted?
4. Do all employees use data to make decision or still they use intuitions? If yes, to what extent?